

Query Match 100.0%; Score 327; DB 6; Length 327;
Best Local Similarity 100.0%; Pred. No. 2.7e-66;
Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGATCTGAATTCGAGCTCGCTGTTGGGCTCGCGGTTGAGGACAAACTCTTCGCGGTCTTTT 60
DB 1 AGATCTGAATTCGAGCTCGCTGTTGGGCTCGCGGTTGAGGACAAACTCTTCGCGGTCTTTT 60

QY 61 CCACTACTCTTTGGATCGGAAACCCCGTCCGCTCCGAAACGGTACTCCGCAACCGAGGACC 120
DB 61 CCACTACTCTTTGGATCGGAAACCCCGTCCGCTCCGAAACGGTACTCCGCAACCGAGGACC 120

QY 121 TGAGCGAGTCCGCATCGACCGGATCGGAAACCTCTCGAAGAGCGCTTAACAGTCAC 180
DB 121 TGAGCGAGTCCGCATCGACCGGATCGGAAACCTCTCGAAGAGCGCTTAACAGTCAC 180

QY 181 AGTCCCAAGGTAGGCTGAGCACCGTGGCGGGCGGAGCGGTTGGGCTGCGGTTGTTTC 240
DB 181 AGTCCCAAGGTAGGCTGAGCACCGTGGCGGGCGGAGCGGTTGGGCTGCGGTTGTTTC 240

QY 241 TGGCGGAGGTGCTGCTGATGATGTAATTAAGTAGGCGGTCTTGAGACGGCGGATGGTCG 300
DB 241 TGGCGGAGGTGCTGCTGATGATGTAATTAAGTAGGCGGTCTTGAGACGGCGGATGGTCG 300

QY 301 AGGTGAGGTGGCAGGCTTGAGATCT 327
DB 301 AGGTGAGGTGGCAGGCTTGAGATCT 327

RESULT 2
BD268204
LOCUS Adenovirus vector, packaging cell line, composition and method for production and use. 7469 bp DNA linear PAT 17-JUL-2003
DEFINITION
ACCESSION BD268204.1 GI:33077972
VERSION BD268204.1
KEYWORDS JP 2002534130-A/8.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 7469)
AUTHORS Nemerow, G.R.; Seggern, D.J.V., Hallenbeck, P.L., Stevenson, S.C. and Skripchenko, Y.
TITLE Adenovirus vector, packaging cell line, composition and method for production and use
JOURNAL Patent: JP 2002534130-A 8 15-OCT-2002;
COMMENT NOVARTIS AG, THE SCRIPPS RESEARCH INSTITUTE
OS Artificial Sequence
PN JP 2002534130-A/8
PD 15-OCT-2002
PF 14-JAN-2000 JP 2000593765
PR 14-JAN-1999 US 60/115920
PI GLEN ROBERT NEMEROW, DANIEL J VON SEGGERN, PAUL L HALLENBECK, PI SUSAN C STEVENSON, YELENA SKRIPCHENKO
PC C12N15/09, A61K35/76, A61K48/00, A61P35/00, A61P43/00, A61P43/00, C12N15/10,
PC C12N7/00, C12Q1/68, G01N33/53, G01N33/566, C12N15/00, C12N5/00 CC
Description of Artificial Sequence: plasmid
FH key Location/Qualifiers
FT source 1..7469
FT /organism='Artificial Sequence'.
FEATURES
source
1..7469
/organism='synthetic construct'
/mol_type='genomic DNA'
/db_xref='taxon:32630'

ORIGIN
Query Match 99.4%; Score 325; DB 6; Length 7469;
Best Local Similarity 100.0%; Pred. No. 7.7e-66;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GATCTGAATTCGAGCTCGCTGTTGGGCTCGCGGTTGAGGACAAACTCTTCGCGGTCTTTT 61

DB 908 GATCTGAATTCGAGCTCGCTGTTGGGCTCGCGGTTGAGGACAAACTCTTCGCGGTCTTTT 967

QY 62 CAGTACTCTTTGGATCGGAAACCCCGTCCGCTCCGAAACGGTACTCCGCCACCGAGGACCT 121
DB 968 CAGTACTCTTTGGATCGGAAACCCCGTCCGCTCCGAAACGGTACTCCGCCACCGAGGACCT 1027

QY 122 GAGCGAGTCCGCATCGACCGGATCGGAAACCTCTCGAAGAGCGCTTAACAGTCACA 181
DB 1028 GAGCGAGTCCGCATCGACCGGATCGGAAACCTCTCGAAGAGCGCTTAACAGTCACA 1087

QY 182 GTCCCAAGGTAGGCTGAGCACCGTGGCGGGCGGAGCGGTTGGGCTGCGGTTGTTTCT 241
DB 1088 GTCCCAAGGTAGGCTGAGCACCGTGGCGGGCGGAGCGGTTGGGCTGCGGTTGTTTCT 1147

QY 242 GCGCGAGGTGCTGCTGATGATGTAATTAAGTAGGCGGTCTTGAGACGGCGGATGGTCGA 301
DB 1148 GCGCGAGGTGCTGCTGATGATGTAATTAAGTAGGCGGTCTTGAGACGGCGGATGGTCGA 1207

QY 302 GGTGAGGTGGCAGGCTTGAGATC 326
DB 1208 GGTGAGGTGGCAGGCTTGAGATC 1232

RESULT 3
AX356037
LOCUS Sequence 8 from Patent WO0183729. 7469 bp DNA linear PAT 06-FEB-2002
DEFINITION
ACCESSION AX356037
VERSION AX356037.1 GI:18620599
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Nemerow, G.R., von Seggern, D.J. and Friedlander, M.
TITLE Vectors for ocular transduction and use thereof for genetic therapy
JOURNAL Patent: WO 0183729-A 8 08-NOV-2001;
COMMENT Novartis AG (CH); The Scripps Research Institute (US); Nemerow, Glen R. (US); Von Seggern, Daniel J. (US); Friedlander, Marty (US)
FEATURES
source
1..7469
/organism='synthetic construct'
/mol_type='unassigned DNA'
/db_xref='taxon:32630'
/note='plasmid'

ORIGIN
Query Match 99.4%; Score 325; DB 6; Length 7469;
Best Local Similarity 100.0%; Pred. No. 7.7e-66;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GATCTGAATTCGAGCTCGCTGTTGGGCTCGCGGTTGAGGACAAACTCTTCGCGGTCTTTT 61
DB 908 GATCTGAATTCGAGCTCGCTGTTGGGCTCGCGGTTGAGGACAAACTCTTCGCGGTCTTTT 967

QY 62 CAGTACTCTTTGGATCGGAAACCCCGTCCGCTCCGAAACGGTACTCCGCCACCGAGGACCT 121
DB 968 CAGTACTCTTTGGATCGGAAACCCCGTCCGCTCCGAAACGGTACTCCGCCACCGAGGACCT 1027

QY 122 GAGCGAGTCCGCATCGACCGGATCGGAAACCTCTCGAAGAGCGCTTAACAGTCACA 181
DB 1028 GAGCGAGTCCGCATCGACCGGATCGGAAACCTCTCGAAGAGCGCTTAACAGTCACA 1087

QY 182 GTCCCAAGGTAGGCTGAGCACCGTGGCGGGCGGAGCGGTTGGGCTGCGGTTGTTTCT 241
DB 1088 GTCCCAAGGTAGGCTGAGCACCGTGGCGGGCGGAGCGGTTGGGCTGCGGTTGTTTCT 1147

QY 242 GCGCGAGGTGCTGCTGATGATGTAATTAAGTAGGCGGTCTTGAGACGGCGGATGGTCGA 301
DB 1148 GCGCGAGGTGCTGCTGATGATGTAATTAAGTAGGCGGTCTTGAGACGGCGGATGGTCGA 1207

QY 302 GGTGAGGTGTGGCAGGCTTGAGATC 326
|||||
Db 1208 GGTGAGGTGTGGCAGGCTTGAGATC 1232

RESULT 4
BD021936 7469 bp DNA linear PAT 27-AUG-2002
LOCUS
DEFINITION Packaging cell systems for use in promotion of the development of high-capacity adenoviral vectors.
ACCESSION BD021936
VERSION BD021936.1 GI:22563159
KEYWORDS JP 2001505047-A/8.
SOURCE unidentified
ORGANISM unclassified.

REFERENCE 1 (bases 1 to 7469)
AUTHORS Memrow,G.R. and Seggern,D.J.V.
TITLE Packaging cell systems for use in promotion of the development of high-capacity adenoviral vectors
JOURNAL Patent: JP 2001505047-A 8 17-APR-2001;
NOVARTIS AG,THE SCRIPPS RESEARCH INSTITUTE
COMMENT PN JP 2001505047-A/8
PD 17-APR-2001
PF 24-SEP-1997 JP 1998515273
PR 25-SEP-1996 US 08/719806
PI GLEN R MEMROW DANIEL J VON SEGGERN
PC C12N5/10,C07K14/075,C12N15/09/A61K31/711,A61K35/76,A61K48/00,
PC A61P35/00,
PC C12N5/00,C12N15/00
PC C12N5/00,C12N15/00
CC Strandedness: Double;
CC Topology: Circular;
FH Key Location/Qualifiers.

FEATURES
source
1..7469
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 99.4%; Score 325; DB 6; Length 7469;
Best Local Similarity 100.0%; Pred. No. 7.7e-66;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GATCTGAATTCGAGCTCGCTGTGGCTCGCGTTGAGACAACTCTTCGGGCTTTTC 61
|||||
Db 908 GATCTGAATTCGAGCTCGCTGTGGCTCGCGTTGAGACAACTCTTCGGGCTTTTC 967

QY 62 CAGTACTCTTGATCGGAACCCGTCGGCTCCGAACCGGTACTCCGCCACCGAGGACCT 121
|||||
Db 968 CAGTACTCTTGATCGGAACCCGTCGGCTCCGAACCGGTACTCCGCCACCGAGGACCT 1027

QY 122 GAGCAGTCCGATCGACCGGATCGGAAACCTCTCGAAGAAAGGCGTCTAACCACTACA 181
|||||
Db 1028 GAGCAGTCCGATCGACCGGATCGGAAACCTCTCGAAGAAAGGCGTCTAACCACTACA 1087

QY 182 GTCGAAGGTAGCTGAGCACCCTGCGGGCGGCGAGCGGTGCGGGTCTTTTCT 241
|||||
Db 1088 GTCGAAGGTAGCTGAGCACCCTGCGGGCGGCGAGCGGTGCGGGTCTTTTCT 1147

QY 242 GCGGAGGTGCTGATGATGATTAATTAAGTAGCGGTCTTGAGACGCGGATGCTGA 301
|||||
Db 1148 GCGGAGGTGCTGATGATGATTAATTAAGTAGCGGTCTTGAGACGCGGATGCTGA 1207

QY 302 GGTGAGGTGTGGCAGGCTTGAGATC 326
|||||
Db 1208 GGTGAGGTGTGGCAGGCTTGAGATC 1232

RESULT 5
BD268212 10610 bp DNA linear PAT 17-JUL-2003
LOCUS
DEFINITION Adenovirus vector, packaging cell line, composition and method for

production and use.
BD268212
BD268212.1 GI:33077980
JP 2002534130-A/16.
synthetic construct
other sequences; artificial sequences.
1 (bases 1 to 10610)
Nemerow,G.R., Seggern,D.J.V., Hallenbeck,P.L., Stevenson,S.C. and Skripchenko,Y.
Adenovirus vector, packaging cell line, composition and method for production and use
Patent: JP 2002534130-A 16 15-OCT-2002;
NOVARTIS AG,THE SCRIPPS RESEARCH INSTITUTE
OS Artificial Sequence
PN JP 2002534130-A/16
PD 15-OCT-2002
PF 14-JAN-2000 JP 2000593765
PR 14-JAN-1999 US 60/115920
PI GLEN ROBERT NEMEROW,DANIEL J VON SEGGERN,PAUL L HALLENBECK, PI SUSAN C STEVENSON,YELENA SKRIPCHENKO
PC C12N15/09,A61K35/76,A61K48/00,A61P35/00,A61P43/00,
PC C12N5/10,
PC C12N5/00,C12O1/68,G01N33/53,G01N33/566,C12N15/00,C12N5/00 CC
Description of Artificial Sequence: plasmid
FH Key Location/Qualifiers
FT source 1..10610
/organism="Artificial Sequence".
Location/Qualifiers
1..10610
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

ORIGIN
Query Match 99.4%; Score 325; DB 6; Length 10610;
Best Local Similarity 100.0%; Pred. No. 7.6e-66;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GATCTGAATTCGAGCTCGCTGTGGCTCGCGTTGAGACAACTCTTCGGGCTTTTC 61
|||||
Db 4049 GATCTGAATTCGAGCTCGCTGTGGCTCGCGTTGAGACAACTCTTCGGGCTTTTC 4108

QY 62 CAGTACTCTTGATCGGAACCCGTCGGCTCCGAACCGGTACTCCGCCACCGAGGACCT 121
|||||
Db 4109 CAGTACTCTTGATCGGAACCCGTCGGCTCCGAACCGGTACTCCGCCACCGAGGACCT 4168

QY 122 GAGCAGTCCGATCGACCGGATCGGAAACCTCTCGAAGAAAGGCGTCTAACCACTACA 181
|||||
Db 4169 GAGCAGTCCGATCGACCGGATCGGAAACCTCTCGAAGAAAGGCGTCTAACCACTACA 4228

QY 182 GTCGAAGGTAGCTGAGCACCCTGCGGGCGGCGAGCGGTGCGGGTCTTTTCT 241
|||||
Db 4229 GTCGAAGGTAGCTGAGCACCCTGCGGGCGGCGAGCGGTGCGGGTCTTTTCT 4288

QY 242 GCGGAGGTGCTGATGATGATTAATTAAGTAGCGGTCTTGAGACGCGGATGCTGA 301
|||||
Db 4289 GCGGAGGTGCTGATGATGATTAATTAAGTAGCGGTCTTGAGACGCGGATGCTGA 4348

QY 302 GGTGAGGTGTGGCAGGCTTGAGATC 326
|||||
Db 4349 GGTGAGGTGTGGCAGGCTTGAGATC 4373

RESULT 6
AX356045 10610 bp DNA linear PAT 06-FEB-2002
LOCUS
DEFINITION Sequence 16 from Patent WO0183729.
ACCESSION AX356045
VERSION AX356045.1 GI:18620607
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM synthetic construct

```
other sequences; artificial sequences.
1
REFERENCE
  AUTHORS Nemerow,G.R., von Seggern,D.J. and Friedlander,M.
  TITLE Vectors for ocular transduction and use thereof for genetic therapy
  JOURNAL Patent: WO 0183729-A 16 08-NOV-2001;
  Novartis AG (CH) ; The Scripps Research Institute (US) ; Nemerow,
  Glen R. (US) ; Von Seggern, Daniel J. (US) ; Friedlander, Marty
  (US)
FEATURES
  source Location/Qualifiers
    1..10610
      /organism="synthetic construct"
      /mol_type="unassigned DNA"
      /db_xref="taxon:32630"
      /note="plasmid"
ORIGIN
  Query Match 99.4%; Score 325; DB 6; Length 10610;
  Best Local Similarity 100.0%; Pred. No. 7.6e-66;
  Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  QY 2 GATCTGAATTCGAGCTCGCTGTTGGCTCGCGGTTGAGGACAACTCTTCGCGGTCTTTC 61
  Db 4049 GATCTGAATTCGAGCTCGCTGTTGGCTCGCGGTTGAGGACAACTCTTCGCGGTCTTTC 4108
  QY 62 CAGTACTCTTGATCGGAACCCGTCGGCTCCGAAACGGTACTCCGCCACCGAGGACCT 121
  Db 4109 CAGTACTCTTGATCGGAACCCGTCGGCTCCGAAACGGTACTCCGCCACCGAGGACCT 4168
  QY 122 GAGCGAGTCCGATCGACCGGATCGGAACCTCTCGAAGAGGGTCTAACCACTCACA 181
  Db 4169 GAGCGAGTCCGATCGACCGGATCGGAACCTCTCGAAGAGGGTCTAACCACTCACA 4228
  QY 182 GTCCAAAGTAGGCTGAGCACCCTGCGCGGCGAGCGGGTGGCGGTGGCGGTGGTCT 241
  Db 4229 GTCCAAAGTAGGCTGAGCACCCTGCGCGGCGAGCGGGTGGCGGTGGCGGTGGTCT 4288
  QY 242 GCGGAGGTGCTGCTGATGATGTAATTAAGTAGGCGGTCTTGAGACGGCGATGGTGA 301
  Db 4289 GCGGAGGTGCTGCTGATGATGTAATTAAGTAGGCGGTCTTGAGACGGCGATGGTGA 4348
  QY 302 GGTGAGGTGTGGCAGGCTTGAGATC 326
  Db 4349 GGTGAGGTGTGGCAGGCTTGAGATC 4373
RESULT 7
BD021944 10610 bp DNA linear PAT 27-AUG-2002
LOCUS Packaging cell systems for use in promotion of the development of
DEFINITION high-capacity adenoviral vectors.
ACCESSION BD021944
VERSION BD021944.1 GI:22563167
KEYWORDS JP 2001505047-A/16.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 10610)
AUTHORS Nemerow,G.R. and Seggern,D.J.V.
TITLE Packaging cell systems for use in promotion of the development of
JOURNAL high-capacity adenoviral vectors
COMMENT Patent: JP 2001505047-A 16 17-APR-2001;
NOVARTIS AG,THE SCRIPPS RESEARCH INSTITUTE
PN JP 2001505047-A/16
PF 17-APR-2001
PR 24-SEP-1997 JP 1998515273
PR 25-SEP-1996 US 08/719806
PI GLEN R NEMEROW,DANIEL J VON SEGGERN
PC C12N5/10,C07K14/075,C12N15/09//A61K31/711.A61K35/76,A61K48/00,
PC A61P35/00,
PC C12N5/00,C12N15/00
CC Strandedness: Double;
CC Topology: Circular;
CC Key Location/Qualifiers.
PH Key

other sequences; artificial sequences.
1
REFERENCE
  AUTHORS Nemerow,G.R., von Seggern,D.J. and Friedlander,M.
  TITLE Vectors for ocular transduction and use thereof for genetic therapy
  JOURNAL Patent: WO 0183729-A 16 08-NOV-2001;
  Novartis AG (CH) ; The Scripps Research Institute (US) ; Nemerow,
  Glen R. (US) ; Von Seggern, Daniel J. (US) ; Friedlander, Marty
  (US)
FEATURES
  source Location/Qualifiers
    1..10610
      /organism="synthetic construct"
      /mol_type="unassigned DNA"
      /db_xref="taxon:32630"
      /note="plasmid"
ORIGIN
  Query Match 99.4%; Score 325; DB 6; Length 10610;
  Best Local Similarity 100.0%; Pred. No. 7.6e-66;
  Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  QY 2 GATCTGAATTCGAGCTCGCTGTTGGCTCGCGGTTGAGGACAACTCTTCGCGGTCTTTC 61
  Db 4049 GATCTGAATTCGAGCTCGCTGTTGGCTCGCGGTTGAGGACAACTCTTCGCGGTCTTTC 4108
  QY 62 CAGTACTCTTGATCGGAACCCGTCGGCTCCGAAACGGTACTCCGCCACCGAGGACCT 121
  Db 4109 CAGTACTCTTGATCGGAACCCGTCGGCTCCGAAACGGTACTCCGCCACCGAGGACCT 4168
  QY 122 GAGCGAGTCCGATCGACCGGATCGGAACCTCTCGAAGAGGGTCTAACCACTCACA 181
  Db 4169 GAGCGAGTCCGATCGACCGGATCGGAACCTCTCGAAGAGGGTCTAACCACTCACA 4228
  QY 182 GTCCAAAGTAGGCTGAGCACCCTGCGCGGCGAGCGGGTGGCGGTGGCGGTGGTCT 241
  Db 4229 GTCCAAAGTAGGCTGAGCACCCTGCGCGGCGAGCGGGTGGCGGTGGCGGTGGTCT 4288
  QY 242 GCGGAGGTGCTGCTGATGATGTAATTAAGTAGGCGGTCTTGAGACGGCGATGGTGA 301
  Db 4289 GCGGAGGTGCTGCTGATGATGTAATTAAGTAGGCGGTCTTGAGACGGCGATGGTGA 4348
  QY 302 GGTGAGGTGTGGCAGGCTTGAGATC 326
  Db 4349 GGTGAGGTGTGGCAGGCTTGAGATC 4373
RESULT 8
BD268211/c 14455 bp DNA linear PAT 17-JUL-2003
LOCUS Adenovirus vector, packaging cell line, composition and method for
DEFINITION production and use.
ACCESSION BD268211
VERSION BD268211.1 GI:33077979
KEYWORDS JP 2002534130-A/15.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 14455)
AUTHORS Nemerow,G.R., Seggern,D.J.V., Hallenbeck,P.L., Stevenson,S.C. and
Skripchenko,Y.
TITLE Adenovirus vector, packaging cell line, composition and method for
JOURNAL production and use
COMMENT Patent: JP 2002534130-A 15 15-OCT-2002;
NOVARTIS AG,THE SCRIPPS RESEARCH INSTITUTE
OS Artificial Sequence
PN JP 2002534130-A/15
PD 15-OCT-2002
PF 14-JAN-2000 JP 2000593765
PF 14-JAN-1999 US 60/115920
PI GLEN ROBERT NEMEROW,DANIEL J VON SEGGERN,PAUL L HALLENBECK, PI
SUSAN C STEVENSON,YELENA SKRIPCHENKO
PC C12N15/09,A61K35/76,A61K48/00,A61P35/00,A61P43/00,A61P43/00,
PC C12N5/10,
PC C12N7/00,C12Q1/68,G01N33/53,G01N33/566,C12N15/00,C12N5/00 CC
Description of Artificial Sequence: plasmid
FH Key Location/Qualifiers
FT source 1..14455 /organism='Artificial Sequence'.
FT source Location/Qualifiers
  1..14455
    /organism="synthetic construct"
    /mol_type="genomic DNA"
    /db_xref="taxon:32630"
```

```
ORIGIN
Query Match          99.4%; Score 325; DB 6; Length 14455;
Best Local Similarity 100.0%; Pred. No. 7.6e-66;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GATCTGAATTCGAGTCGCTGTTGGCTCGCGGTTGAGGACAACTCTTCGCGGTCCTTTC 61
DB 13315 GATCTGAATTCGAGTCGCTGTTGGCTCGCGGTTGAGGACAACTCTTCGCGGTCCTTTC 13256

QY 62 CAGTACTCTTGATCGGAAACCCGTCGCGCTCCGAAACGGTACTCCGCCACCGAGGACCT 121
DB 13255 CAGTACTCTTGATCGGAAACCCGTCGCGCTCCGAAACGGTACTCCGCCACCGAGGACCT 13196

QY 122 GAGCAGTCCGCATCGACCGGATCGGAAACCTCTCGAGAAAGCGCTTAACCAAGTCA 181
DB 13195 GAGCAGTCCGCATCGACCGGATCGGAAACCTCTCGAGAAAGCGCTTAACCAAGTCA 13136

QY 182 GTCCCAAGGTAGGCTGAGCACCCTGCGCGGCGGAGCGGTCGCGGTCGGGTTGTTTCT 241
DB 13135 GTCCCAAGGTAGGCTGAGCACCCTGCGCGGCGGAGCGGTCGCGGTCGGGTTGTTTCT 13076

QY 242 GCGGAGGTCGCTGATGATGATTAATTAAGTAGCGGCTCTTGAGACGGCGATGTCGA 301
DB 13075 GCGGAGGTCGCTGATGATGATTAATTAAGTAGCGGCTCTTGAGACGGCGATGTCGA 13016

QY 302 GGTGAGGTGTGGCAGGCTTGAGATC 326
DB 13015 GGTGAGGTGTGGCAGGCTTGAGATC 12991

RESULT 9
LOCUS AX356044/c 14455 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 15 from Patent WO0183729.
ACCESSION AX356044
VERSION AX356044.1 GI:18620606
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE
1 Nemerow,G.R., von Seggern,D.J. and Friedlander,M.
  Vectors for ocular transduction and use thereof for genetic therapy
  Patent: WO 0183729-A 15 08-NOV-2001;
  Novartis AG (CH); The Scripps Research Institute (US); Nemerow,
  Glen R. (US); Von Seggern, Daniel J. (US); Friedlander, Marty
  (US)
FEATURES
source
  Location/Qualifiers
  1. 14455
    /organism="synthetic construct"
    /mol_type="unassigned DNA"
    /db_xref="taxon:32630"
    /note="plasmid"
ORIGIN
Query Match          99.4%; Score 325; DB 6; Length 14455;
Best Local Similarity 100.0%; Pred. No. 7.6e-66;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GATCTGAATTCGAGTCGCTGTTGGCTCGCGGTTGAGGACAACTCTTCGCGGTCCTTTC 61
DB 13315 GATCTGAATTCGAGTCGCTGTTGGCTCGCGGTTGAGGACAACTCTTCGCGGTCCTTTC 13256

QY 62 CAGTACTCTTGATCGGAAACCCGTCGCGCTCCGAAACGGTACTCCGCCACCGAGGACCT 121
DB 13255 CAGTACTCTTGATCGGAAACCCGTCGCGCTCCGAAACGGTACTCCGCCACCGAGGACCT 13196

QY 122 GAGCAGTCCGCATCGACCGGATCGGAAACCTCTCGAGAAAGCGCTTAACCAAGTCA 181
DB 13195 GAGCAGTCCGCATCGACCGGATCGGAAACCTCTCGAGAAAGCGCTTAACCAAGTCA 13136

QY 182 GTCCCAAGGTAGGCTGAGCACCCTGCGCGGCGGAGCGGTCGCGGTCGGGTTGTTTCT 241
DB 13135 GTCCCAAGGTAGGCTGAGCACCCTGCGCGGCGGAGCGGTCGCGGTCGGGTTGTTTCT 13076

QY 242 GCGGAGGTCGCTGATGATGATTAATTAAGTAGCGGCTCTTGAGACGGCGATGTCGA 301
DB 13075 GCGGAGGTCGCTGATGATGATTAATTAAGTAGCGGCTCTTGAGACGGCGATGTCGA 13016

QY 302 GGTGAGGTGTGGCAGGCTTGAGATC 326
DB 13015 GGTGAGGTGTGGCAGGCTTGAGATC 12991

RESULT 9
LOCUS AX356044/c 14455 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 15 from Patent WO0183729.
ACCESSION AX356044
VERSION AX356044.1 GI:18620606
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE
1 Nemerow,G.R., von Seggern,D.J. and Friedlander,M.
  Vectors for ocular transduction and use thereof for genetic therapy
  Patent: WO 0183729-A 15 08-NOV-2001;
  Novartis AG (CH); The Scripps Research Institute (US); Nemerow,
  Glen R. (US); Von Seggern, Daniel J. (US); Friedlander, Marty
  (US)
FEATURES
source
  Location/Qualifiers
  1. 14455
    /organism="synthetic construct"
    /mol_type="unassigned DNA"
    /db_xref="taxon:32630"
    /note="plasmid"
ORIGIN
Query Match          99.4%; Score 325; DB 6; Length 14455;
Best Local Similarity 100.0%; Pred. No. 7.6e-66;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GATCTGAATTCGAGTCGCTGTTGGCTCGCGGTTGAGGACAACTCTTCGCGGTCCTTTC 61
DB 13315 GATCTGAATTCGAGTCGCTGTTGGCTCGCGGTTGAGGACAACTCTTCGCGGTCCTTTC 13256

QY 62 CAGTACTCTTGATCGGAAACCCGTCGCGCTCCGAAACGGTACTCCGCCACCGAGGACCT 121
DB 13255 CAGTACTCTTGATCGGAAACCCGTCGCGCTCCGAAACGGTACTCCGCCACCGAGGACCT 13196

QY 122 GAGCAGTCCGCATCGACCGGATCGGAAACCTCTCGAGAAAGCGCTTAACCAAGTCA 181
DB 13195 GAGCAGTCCGCATCGACCGGATCGGAAACCTCTCGAGAAAGCGCTTAACCAAGTCA 13136

QY 182 GTCCCAAGGTAGGCTGAGCACCCTGCGCGGCGGAGCGGTCGCGGTCGGGTTGTTTCT 241
DB 13135 GTCCCAAGGTAGGCTGAGCACCCTGCGCGGCGGAGCGGTCGCGGTCGGGTTGTTTCT 13076

QY 242 GCGGAGGTCGCTGATGATGATTAATTAAGTAGCGGCTCTTGAGACGGCGATGTCGA 301
DB 13075 GCGGAGGTCGCTGATGATGATTAATTAAGTAGCGGCTCTTGAGACGGCGATGTCGA 13016

QY 302 GGTGAGGTGTGGCAGGCTTGAGATC 326
DB 13015 GGTGAGGTGTGGCAGGCTTGAGATC 12991

RESULT 9
LOCUS AX356044/c 14455 bp DNA linear PAT 27-AUG-2002
DEFINITION Packaging cell systems for use in promotion of the development of
  high-capacity adenoviral vectors.
ACCESSION BD021943
VERSION BD021943.1 GI:22563166
KEYWORDS JP 2001505047-A/15.
SOURCE unidentified
ORGANISM unidentified
unclassified.
REFERENCE
1 (bases 1 to 14455)
  Memerow,G.R. and Seggern,D.J.V.
  Packaging cell systems for use in promotion of the development of
  high-capacity adenoviral vectors
  Patent: JP 2001505047-A 15 17-APR-2001;
  NOVARTIS AG,THE SCRIPPS RESEARCH INSTITUTE
  PN JP 2001505047-A/15
  PD 17-APR-2001
  PF 24-SEP-1997 JP 1998515273
  PI GLEN R MEMEROW DANIEL J VON SEGGERN
  PC C12N5/10,C07K14/075,C12N15/09//A61K31/711,A61K35/76,A61K48/00,
  PC A61P35/00,
  PC C12N5/00,C12N15/00
  CC Strandedness: Double;
  CC Topology: Circular;
  FH Key Location/Qualifiers
FEATURES
source
  Location/Qualifiers
  1. 14455
    /organism="unidentified"
    /mol_type="genomic DNA"
    /db_xref="taxon:32644"
ORIGIN
Query Match          99.4%; Score 325; DB 6; Length 14455;
Best Local Similarity 100.0%; Pred. No. 7.6e-66;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GATCTGAATTCGAGTCGCTGTTGGCTCGCGGTTGAGGACAACTCTTCGCGGTCCTTTC 61
DB 13315 GATCTGAATTCGAGTCGCTGTTGGCTCGCGGTTGAGGACAACTCTTCGCGGTCCTTTC 13256

QY 62 CAGTACTCTTGATCGGAAACCCGTCGCGCTCCGAAACGGTACTCCGCCACCGAGGACCT 121
DB 13255 CAGTACTCTTGATCGGAAACCCGTCGCGCTCCGAAACGGTACTCCGCCACCGAGGACCT 13196

QY 122 GAGCAGTCCGCATCGACCGGATCGGAAACCTCTCGAGAAAGCGCTTAACCAAGTCA 181
DB 13195 GAGCAGTCCGCATCGACCGGATCGGAAACCTCTCGAGAAAGCGCTTAACCAAGTCA 13136

QY 182 GTCCCAAGGTAGGCTGAGCACCCTGCGCGGCGGAGCGGTCGCGGTCGGGTTGTTTCT 241
DB 13135 GTCCCAAGGTAGGCTGAGCACCCTGCGCGGCGGAGCGGTCGCGGTCGGGTTGTTTCT 13076

QY 242 GCGGAGGTCGCTGATGATGATTAATTAAGTAGCGGCTCTTGAGACGGCGATGTCGA 301
DB 13075 GCGGAGGTCGCTGATGATGATTAATTAAGTAGCGGCTCTTGAGACGGCGATGTCGA 13016

QY 302 GGTGAGGTGTGGCAGGCTTGAGATC 326
DB 13015 GGTGAGGTGTGGCAGGCTTGAGATC 12991
```

```
Db 13015 GGTGAGGTGTGGCAGGCTTGAGATC 12991

RESULT 11
LOCUS I09267 5365 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 34 from Patent WO 8901940.
ACCESSION I09267
VERSION I09267.1 GI:588051
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 5365)
AUTHORS Fisher,R.A., Gilbert,W., Sato,V.L., Flavell,R.A., Maraganore,J.M.
and Liu,T.R.
TITLE DNA SEQUENCES, RECOMBINANT DNA MOLECULES AND PROCESSES FOR
PRODUCING SOLUBLE T4 PROTEINS
JOURNAL Patent: WO 8901940-A 34 09-MAR-1989;
FEATURES
source
Location/Qualifiers
1..5365
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 94.6%; Score 309.2; DB 6; Length 5365;
Best Local Similarity 99.0%; Pred. No. 3.9e-62;
Matches 311; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 13 GAGCTCGCTGTGGGCTCGCGGTGAGGACAAACTCTTCGCGGTCTTTCCAGTACTCTTG 72
Db 611 GGGCCAGCTGTGGGCTCGCGGTGAGGACAAACTCTTCGCGGTCTTTCCAGTACTCTTG 670

Qy 73 GATCGAAACCCGTCGGCTCCGAAACGGTACTCTCGCCACCGAGGACCTGAGCGAGTCCG 132
Db 671 GATCGAAACCCGTCGGCTCCGAAACGGTACTCTCGCCACCGAGGACCTGAGCGAGTCCG 730

Qy 133 CATCACCGGATCGGAAACCTCTCGAGAAAGCGTCTAACCACTCAGTCACAGTCGCAAGTA 192
Db 731 CATCACCGGATCGGAAACCTCTCGAGAAAGCGTCTAACCACTCAGTCACAGTCGCAAGTA 790

Qy 193 GGCTGAGCACCGTCGGCGGCGGACGGGTGCGGGTCTGTTTCTGCGGAGGTGC 252
Db 791 GGCTGAGCACCGTCGGCGGCGGACGGGTGCGGGTCTGTTTCTGCGGAGGTGC 850

Qy 253 TGCTGATGATGTAATTAAGTAGCGGTCTTGAGACGGCGGATGTCGAGGTGAGGTGTG 312
Db 851 TGCTGATGATGTAATTAAGTAGCGGTCTTGAGACGGCGGATGTCGAGGTGAGGTGTG 910

Qy 313 GCAGGCTTGAGATC 326
Db 911 GCAGGCTTGAGATC 924

RESULT 12
LOCUS I09270 5413 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 37 from Patent WO 8901940.
ACCESSION I09270
VERSION I09270.1 GI:588054
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 5413)
AUTHORS Fisher,R.A., Gilbert,W., Sato,V.L., Flavell,R.A., Maraganore,J.M.
and Liu,T.R.
TITLE DNA SEQUENCES, RECOMBINANT DNA MOLECULES AND PROCESSES FOR
PRODUCING SOLUBLE T4 PROTEINS
JOURNAL Patent: WO 8901940-A 37 09-MAR-1989;
FEATURES
source
Location/Qualifiers
1..5413
/organism="unknown"

ORIGIN
Query Match 94.6%; Score 309.2; DB 6; Length 5518;
Best Local Similarity 99.0%; Pred. No. 3.9e-62;
Matches 311; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 13 GAGCTCGCTGTGGGCTCGCGGTGAGGACAAACTCTTCGCGGTCTTTCCAGTACTCTTG 72
Db 611 GGGCCAGCTGTGGGCTCGCGGTGAGGACAAACTCTTCGCGGTCTTTCCAGTACTCTTG 670

Qy 73 GATCGAAACCCGTCGGCTCCGAAACGGTACTCTCGCCACCGAGGACCTGAGCGAGTCCG 132
Db 671 GATCGAAACCCGTCGGCTCCGAAACGGTACTCTCGCCACCGAGGACCTGAGCGAGTCCG 730

Qy 133 CATCACCGGATCGGAAACCTCTCGAGAAAGCGTCTAACCACTCAGTCACAGTCGCAAGTA 192
Db 731 CATCACCGGATCGGAAACCTCTCGAGAAAGCGTCTAACCACTCAGTCACAGTCGCAAGTA 790

Qy 193 GGCTGAGCACCGTCGGCGGCGGACGGGTGCGGGTCTGTTTCTGCGGAGGTGC 252
Db 791 GGCTGAGCACCGTCGGCGGCGGACGGGTGCGGGTCTGTTTCTGCGGAGGTGC 850

Qy 253 TGCTGATGATGTAATTAAGTAGCGGTCTTGAGACGGCGGATGTCGAGGTGAGGTGTG 312
Db 851 TGCTGATGATGTAATTAAGTAGCGGTCTTGAGACGGCGGATGTCGAGGTGAGGTGTG 910

Qy 313 GCAGGCTTGAGATC 326
Db 911 GCAGGCTTGAGATC 924

RESULT 13
LOCUS I09268 5518 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 35 from Patent WO 8901940.
ACCESSION I09268
VERSION I09268.1 GI:588052
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 5518)
AUTHORS Fisher,R.A., Gilbert,W., Sato,V.L., Flavell,R.A., Maraganore,J.M.
and Liu,T.R.
TITLE DNA SEQUENCES, RECOMBINANT DNA MOLECULES AND PROCESSES FOR
PRODUCING SOLUBLE T4 PROTEINS
JOURNAL Patent: WO 8901940-A 35 09-MAR-1989;
FEATURES
source
Location/Qualifiers
1..5518
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 94.6%; Score 309.2; DB 6; Length 5518;
Best Local Similarity 99.0%; Pred. No. 3.9e-62;
Matches 311; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 13 GAGCTCGCTGTGGGCTCGCGGTGAGGACAAACTCTTCGCGGTCTTTCCAGTACTCTTG 72
Db 611 GGGCCAGCTGTGGGCTCGCGGTGAGGACAAACTCTTCGCGGTCTTTCCAGTACTCTTG 670

Qy 73 GATCGAAACCCGTCGGCTCCGAAACGGTACTCTCGCCACCGAGGACCTGAGCGAGTCCG 132
Db 671 GATCGAAACCCGTCGGCTCCGAAACGGTACTCTCGCCACCGAGGACCTGAGCGAGTCCG 730

Qy 133 CATCACCGGATCGGAAACCTCTCGAGAAAGCGTCTAACCACTCAGTCACAGTCGCAAGTA 192
Db 731 CATCACCGGATCGGAAACCTCTCGAGAAAGCGTCTAACCACTCAGTCACAGTCGCAAGTA 790

Qy 193 GGCTGAGCACCGTCGGCGGCGGACGGGTGCGGGTCTGTTTCTGCGGAGGTGC 252
Db 791 GGCTGAGCACCGTCGGCGGCGGACGGGTGCGGGTCTGTTTCTGCGGAGGTGC 850

Qy 253 TGCTGATGATGTAATTAAGTAGCGGTCTTGAGACGGCGGATGTCGAGGTGAGGTGTG 312
Db 851 TGCTGATGATGTAATTAAGTAGCGGTCTTGAGACGGCGGATGTCGAGGTGAGGTGTG 910

Qy 313 GCAGGCTTGAGATC 326
Db 911 GCAGGCTTGAGATC 924
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:35:42 ; Search time 461.032 Seconds
(without alignments)
4198.742 Million cell updates/sec

Title: US-09-482-682-26

Perfect score: 327

Sequence: 1 agatctgaattcgactgcg.....gtgtggcaggcttgatctc 327

Scoring table: IDENTITY_NUC

Gapop 10_0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	327	100.0	327	3	AA59054 Nucleotid
2	327	100.0	327	6	ABA94266
3	327	100.0	327	10	ADB75112 Adenoviru
4	327	100.0	327	10	ADF48736
5	325	99.4	7469	2	AAV32372 Complete
6	325	99.4	7469	3	AA59043 Nucleotid
7	325	99.4	7469	6	ABA94252 Nucleotid
8	325	99.4	7469	10	ADB75149
9	325	99.4	7469	10	ADF48794 Fibre exp
10	325	99.4	10610	2	AAV32375 Complete
11	325	99.4	10610	3	AA59051 Nucleotid
12	325	99.4	10610	6	ABA94260 Nucleotid
13	325	99.4	10610	10	ADB75157 Plasmid p
14	325	99.4	10610	10	ADF48802 E4/fibre
15	325	99.4	14455	2	AAV32374 Complete
16	325	99.4	14455	3	AA59050 Nucleotid
17	325	99.4	14455	6	ABA94259 Nucleotid
18	325	99.4	14455	10	ADB75156 Plasmid p
19	325	99.4	14455	10	ADF48801 E1/fibre
20	310.2	94.9	7316	2	AAQ04555 Plasmid p

ALIGNMENTS

RESULT 1

AAA59054

ID AAA59054 standard; DNA; 327 BP.

XX

AC AAA59054;

XX

DT 07-NOV-2000 (first entry)

XX

DE Nucleotide sequence of a tripartite leader sequence.

XX

KW Adenovirus; tripartite leader; adenovirus vector particle; gene delivery;

ss.

OS Mastadenovirus.

XX

PN WO200042208-A1.

XX

PD 20-JUL-2000.

XX

PF 14-JAN-2000; 2000WO-EP000265.

XX

PR 14-JAN-1999; 99US-0115920P.

XX

PA (NOVS) NOVARTIS AG.

PA (NOVS) NOVARTIS-ERFINDUNGEN VERW GES MBH.

XX (SCRI) SCRIPPS RES INST.

PI Nemerow GR, Von Seggern DJ, Hallenbeck PL, Stevenson SC;

PI Skripchenko Y;

XX WPI; 2000-476068/41.

XX

PT New nucleic acid comprising an adenovirus tripartite leader nucleotide

for producing high-capacity and targeted vectors for adenovirus-based

gene therapy.

XX

PS Claim 14; Page 169; 212pp; English.

XX

CC The specification describes a nucleic acid molecule comprising an

adenovirus (AV) tripartite leader (TPL) nucleotide with a sequence

comprising two different TPL exons or three same or different TPL exons.

CC The nucleic acid is used to produce an adenovirus vector particle,

deliver an exogenous gene to a target cell, pseudotype recombinant viral

CC vectors, target an adenovirus vector to a cell, produce a modified
 CC adenovirus, deliver a heterologous gene to an animal and produce a
 CC gutless adenoviral vector particle. The present sequence represents a TPL
 CC sequence, which is used to construct nucleic acid molecules of the
 CC invention

SQ Sequence 327 BP; 63 A; 79 C; 116 G; 69 T; 0 U; 0 Other;
 Query Match 100.0%; Score 327; DB 3; Length 327;
 Best Local Similarity 100.0%; Pred. No. 4.8e-80;
 Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AGATCTGAATTCGAGCTCGCTGTTGGGCTCGGGTTGAGGACAAACTCTTCGGGTCTTT 60
 DB 1 AGATCTGAATTCGAGCTCGCTGTTGGGCTCGGGTTGAGGACAAACTCTTCGGGTCTTT 60
 QY 61 CCAGTACTCTTGGATCGGAAACCCGTCGCGCTCCGAAACGGTACTCCGCCACCGAGGGACC 120
 DB 61 CCAGTACTCTTGGATCGGAAACCCGTCGCGCTCCGAAACGGTACTCCGCCACCGAGGGACC 120
 QY 121 TGAGCGAGTCCGATCGACCGGATCGGAAACCTCTCGAGAAAGCGCTTAACAGTCAC 180
 DB 121 TGAGCGAGTCCGATCGACCGGATCGGAAACCTCTCGAGAAAGCGCTTAACAGTCAC 180
 QY 181 AGTCCAGGTAGGTAGCACCCTGCGCGGCGGAGCGGTGCGCGGTGCGGGTTGTTTC 240
 DB 181 AGTCCAGGTAGGTAGCACCCTGCGCGGCGGAGCGGTGCGCGGTGCGGGTTGTTTC 240
 QY 241 TGGCGGAGTGTCTGTGATGATGTAATTAAGTAGCGGTCTTGAGACGCGGATGTCG 300
 DB 241 TGGCGGAGTGTCTGTGATGATGTAATTAAGTAGCGGTCTTGAGACGCGGATGTCG 300
 QY 301 AGGTGAGGTGTGGCAGGCTTGAGATCT 327
 DB 301 AGGTGAGGTGTGGCAGGCTTGAGATCT 327

RESULT 2
 ID ABA94266
 AC ABA94266;
 XX
 XX 07-AUG-2003 (revised)
 DT 13-MAR-2002 (first entry)
 XX
 DE Adenovirus 5 tripartite leader (TPL) partial nucleotide sequence.
 XX
 KW Adenovirus; inverter terminal repeat sequence; ITRS; ocular disease;
 KW fiber protein; photoreceptor; rhodopsin; stargardt disease gene; STDG1;
 KW ophthalmological; antiinflammatory; antidiabetic; cytostatic;
 KW gene therapy; tripartite leader; TPL; ds.
 XX
 OS Human adenovirus type 5.
 XX
 XX WO200183729-A2.
 PN
 XX
 XX 08-NOV-2001.
 XX
 XX 30-APR-2001; 2001WO-EP004863.
 XX
 XX 01-MAY-2000; 2000US-00562934.
 XX
 XX (NOVS) NOVARTIS AG.
 PA (SCHI) SCRIPPS RES INST.
 PA (NEME/) NEMEROW G R.
 PA (VSEGE/) VON SEGGERN D J.
 PA (FRIE/) FRIEDLANDER M.
 XX
 XX Nemerow GR, Von Seggern DJ, Friedlander M;
 PI
 XX
 XX WPI; 2002-082846/11.
 DR
 XX

PT Polynucleotide for making vectors, useful for treating ocular diseases,
 PT e.g., retinitis pigmentosa, comprises adenovirus inverter terminal repeat
 XX sequences, packaging signal and photoreceptor-specific promoter.
 XX
 PS Example 1; Page 122; 149pp; English.

XX The invention provides an isolated polynucleotide comprising adenovirus
 CC (AV) inverter terminal repeat sequences (ITRS), AV packaging signal
 CC operatively linked to ITRS and a photoreceptor-specific promoter. A
 CC recombinant AV vector (AVV) comprising the polynucleotide is useful for
 CC targeted delivery of a gene product to the eye (especially to the
 CC vitreous cavity), for treating an ocular disease, e.g., retinal
 CC degenerative disease, retinitis pigmentosa, Stargardt's disease, diabetic
 CC retinopathies, retinal vascularizations, and retinoblastoma, of a mammal
 CC preferably human. The AAV comprises a fiber protein that specifically or
 CC selectively binds to receptors that are expressed on cells (preferably
 CC photoreceptors in the eye). Preferably, the recombinant virus comprise a
 CC fiber protein from an adenovirus type D subgroup or is a chimeric protein
 CC containing a portion of the N-terminus of an adenovirus type 2 or type 5
 CC penton, and the therapeutic product is a trophic factor, an anti-
 CC apoptotic factor, gene encoding a rhodopsin protein, a wild-type
 CC stargardt disease gene (STDG1), an anti-cancer agent and a protein that
 CC regulates expression of a photoreceptor specific gene product. The viral
 CC nucleic acid of AAV comprises ITRS and packaging signal derived from AAV
 CC subgroup B or C, especially an AV type 2 or type 5. AAV is also useful
 CC for targeted gene therapy, where the vector comprises an AV type 37 fiber
 CC protein or its portion, and selectively transduces photoreceptors and
 CC delivers a gene product encoded by AAV. The present sequence represents a
 CC adenovirus 5 tripartite leader (TPL) partial nucleotide sequence. (Updated
 CC on 07-AUG-2003 to correct OS field.)
 XX

SQ Sequence 327 BP; 63 A; 79 C; 116 G; 69 T; 0 U; 0 Other;
 Query Match 100.0%; Score 327; DB 6; Length 327;
 Best Local Similarity 100.0%; Pred. No. 4.8e-80;
 Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGATCTGAATTCGAGCTCGCTGTTGGGCTCGGGTTGAGGACAAACTCTTCGGGTCTTT 60
 DB 1 AGATCTGAATTCGAGCTCGCTGTTGGGCTCGGGTTGAGGACAAACTCTTCGGGTCTTT 60
 QY 61 CCAGTACTCTTGGATCGGAAACCCGTCGCGCTCCGAAACGGTACTCCGCCACCGAGGGACC 120
 DB 61 CCAGTACTCTTGGATCGGAAACCCGTCGCGCTCCGAAACGGTACTCCGCCACCGAGGGACC 120
 QY 121 TGAGCGAGTCCGATCGACCGGATCGGAAACCTCTCGAGAAAGCGCTTAACAGTCAC 180
 DB 121 TGAGCGAGTCCGATCGACCGGATCGGAAACCTCTCGAGAAAGCGCTTAACAGTCAC 180
 QY 181 AGTCCAGGTAGGTAGCACCCTGCGCGGCGGAGCGGTGCGCGGTGCGGGTTGTTTC 240
 DB 181 AGTCCAGGTAGGTAGCACCCTGCGCGGCGGAGCGGTGCGCGGTGCGGGTTGTTTC 240
 QY 241 TGGCGGAGTGTCTGTGATGATGTAATTAAGTAGCGGTCTTGAGACGCGGATGTCG 300
 DB 241 TGGCGGAGTGTCTGTGATGATGTAATTAAGTAGCGGTCTTGAGACGCGGATGTCG 300
 QY 301 AGGTGAGGTGTGGCAGGCTTGAGATCT 327
 DB 301 AGGTGAGGTGTGGCAGGCTTGAGATCT 327
 RESULT 3
 ID ADB75112
 XX ADB75112 standard; DNA; 327 BP.
 ID
 XX
 AC ADB75112;
 XX
 XX 04-DEC-2003 (first entry)
 DT
 XX
 XX Adenovirus type 5 tripartite leader sequence #1.
 DE
 XX
 XX ophthalmological; antiinflammatory; antidiabetic; gene therapy;
 KW

adenovirus inverted terminal repeat sequence;
 adenovirus packaging signal; photoreceptor-specific promoter;
 adenovirus type 37; adenovirus type D serotype; adenovirus type 2;
 adenovirus type 5; photoreceptor; trophic factor; anti-apoptotic factor;
 rhodopsin; wild-type Stargardt disease gene; STDG1; anti-cancer agent;
 retinal degenerative disease; retinitis pigmentosa; Stargardt's disease;
 diabetic retinopathy; retinal vascularisation; choroideraemia;
 gyrate atrophy; macular dystrophy; retinoblastoma;
 photoreceptor-restricted transgene expression;
 recombinant adenovirus vector; adenovirus type 5; Ad5;
 tripartite leader sequence; TPL; ds.

Human adenovirus type 5.

US2002193327-A1.

19-DEC-2002.

01-MAY-2001; 2001US-00847101.

01-MAY-2000; 2000US-00562934.

(SCRI) SCRIPPS RES INST.

Nemerow GR, Von Seggern DJ, Friedlander M;

WPI; 2003-657234/62.

Novel nucleic acids comprising adenovirus inverted terminal repeat sequences, adenovirus packaging signals operatively linked to the sequences and photoreceptor-specific promoters, useful for treating retinitis pigmentosa.

Example 1; Page 62; 106pp; English.

The invention describes an isolated nucleic acid (I) comprising adenovirus inverted terminal repeat sequence, an adenovirus packaging signal operatively linked to the sequence, and a photoreceptor-specific promoter. A Recombinant adenovirus vector (II) comprising (I) is useful for targeted delivery of a gene product to the eye of a mammal which involves administering (II) that comprises heterologous DNA encoding the gene product or resulting in expression of the gene product, where the recombinant virus comprises a fibre protein that specifically or selectively binds to receptors that are expressed on cells which are photoreceptors, in the eye. The recombinant virus comprises a fibre protein which is an adenovirus type 37, from an adenovirus type D serotype. The fibre is a chimeric protein containing a sufficient portion of the N-terminus of an adenovirus type 2 or type 5 fibre protein for interaction with an adenovirus type 2 or type 5 penton, and a sufficient selective binding to photoreceptors in the eye of a mammal. The encapsulated nucleic acid comprises a photoreceptor-specific promoter operatively linked to a nucleic acid comprising the therapeutic product which is chosen from trophic factor, anti-apoptotic factor, gene encoding a rhodopsin protein, wild-type Stargardt disease gene (STDG1), an anti-cancer agent and a protein that regulates expression of a photoreceptor-specific gene product. The delivery is effected for treatment of an ocular disease such as retinal degenerative disease e.g., retinitis pigmentosa, Stargardt's disease, diabetic retinopathies, retinal vascularisation, choroideraemia, gyrate atrophy or macular dystrophy or retinoblastoma inherited and acquired retinal and neovascular degenerative diseases. The viral nucleic acid comprises an adenovirus inverted terminal repeat (ITR) sequences, and an adenovirus packaging signal operatively linked to the sequence. The ITRs and packaging signal are derived from an adenovirus serotype B or C, or adenovirus type 2 or 5. The viral nucleic acid further comprises a photoreceptor-specific promoter. (II) includes photoreceptor promoters providing a means not only for specific targeting of expression in these cells, but also for photoreceptor-restricted transgene expression. This sequence represents a TPL (tripartite leader sequence) from the adenovirus type 5 genome, used to enhance the expression of complementing adenoviral proteins.

Sequence 327 BP; 63 A; 79 C; 116 G; 69 T; 0 U; 0 Other;

Query Match 100.0%; Score 327; DB 10; Length 327;
 Best Local Similarity 100.0%; Pred. No. 4.8e-80;
 Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AGATCTGAATTCGAGCTCGCTGTTGGGCTCGGGTTGAGGACAAACTCTTCGGGTTCTTT 60
 DB 1 AGATCTGAATTCGAGCTCGCTGTTGGGCTCGGGTTGAGGACAAACTCTTCGGGTTCTTT 60
 QY 61 CCAGTACTCTTGGATCGGAACCCGCTCGGCCTCCGAAACGGTACTCCGCCACCGAGGACC 120
 DB 61 CCAGTACTCTTGGATCGGAACCCGCTCGGCCTCCGAAACGGTACTCCGCCACCGAGGACC 120
 QY 121 TGAGCGAGTCCGCATCGACCCGATCGGAAACCTCTCGAGAAAGCGTCTAACCAAGTAC 180
 DB 121 TGAGCGAGTCCGCATCGACCCGATCGGAAACCTCTCGAGAAAGCGTCTAACCAAGTAC 180
 QY 181 AGTCCAAAGGTAGGCTGAGACCCGTCGGCGGGCGGAGCGGGTGGCGGTTGTTTC 240
 DB 181 AGTCCAAAGGTAGGCTGAGACCCGTCGGCGGGCGGAGCGGGTGGCGGTTGTTTC 240
 QY 241 TGGCGGAGGTCTGCTGATGATGTAATTAAGTAGGCGGTCTTGAGACGGCGGATGTCG 300
 DB 241 TGGCGGAGGTCTGCTGATGATGTAATTAAGTAGGCGGTCTTGAGACGGCGGATGTCG 300
 QY 301 AGGTGAGGTGTCGACGGCTTCAGATCT 327
 DB 301 AGGTGAGGTGTCGACGGCTTCAGATCT 327

RESULT 4

ADP48736

ID ADP48736 standard; DNA; 327 BP.

XX

AC ADP48736;

XX

DT 12-FEB-2004 (first entry)

DE

XX Adenovirus type 5 partial tripartite leader sequence.

KW cytostatic; anti-HIV; gene therapy; HIV gene expression inhibitor;
 KW HIV gene expression activation; adenovirus tripartite leader; TPL;
 KW gutless adenoviral vector particle;
 KW helper-independent fiberless recombinant adenovirus vector;
 KW packaging cell line; pseudotyping; adenovirus vector; gene therapy;
 KW hereditary disorder; tumour; HIV infection; fibre;
 KW fibre-gene-deleted adenoviruses; hygromycin resistance;
 KW tripartite leader sequence; ds.

OS Human adenovirus type 5.

XX

PN US2003157688-A1.

XX

PD 21-AUG-2003.

XX

PF 14-JAN-2000; 2000US-00482682.

XX

PR 14-JAN-1999; 99US-0115920P.

PR 26-JUN-2000; 2000US-00423783.

XX

PA (VSEG/) VON SEGGERN D J.

PA (NEME/) NEMEROW G R.

PA (HALL/) HALLENBECK P.

PA (STEV/) STEVENSON S.

XX (SKRI/) SKRIPCHENKO Y.

PI Von Seggern DJ, Nemerow GR, Hallenbeck P, Stevenson S;

PI Skripchenko Y;

XX WPI; 2003-843463/78.

XX

PT Novel isolated nucleic acid molecule useful for delivering heterologous gene to human or any animal, or for producing gutless adenoviral vector

PT particle.
 XX Claim 14; SEQ ID NO 26; 157pp; English.
 XX The invention describes an isolated nucleic acid molecule (I) comprising an adenovirus tripartite leader (TPL) nucleotide, the TPL nucleotide sequence comprising a first and second different TPL exons or first, second and third same or different TPL exons, the TPL exons chosen from complete or partial TPL exon 1, complete TPL exon 2 and complete TPL exon 3. (I) is useful for delivering a heterologous gene to a human or any animal, or for producing a gutless adenoviral vector particle. A recombinant adenovirus particle (II) is useful for delivering of an exogenous gene to a target cell which involves contacting the cell with an amount of (II) sufficient to infect the cell. A helper-independent fiberless recombinant adenovirus vector genome (III) is useful for producing an adenovirus vector particle containing (III) which involves packaging a packaging cell line which complements replication and packaging of the genome and (III) which is deficient in expressing sufficient functional fiber protein to support assembly of fiber containing particles and harvesting the particle produced by the cell line. (III) is useful for pseudotyping recombinant viral vectors which involves complementing a missing fiber gene of (III) or helper-packaging fiberless recombinant adenovirus vector genome by expressing in packaging cells a fiber gene from a different adenoviral serotype than the recombinant adenovirus vector. (III) is also useful for specifically targeting an adenovirus vector to a cell of choice. (I) is useful for gene therapy. (II) is useful for treating diseases such as hereditary disorder, and for reducing proliferation of tumour cells in a subject, or to disrupt HIV infection. This sequence represents a partial adenovirus tripartite leader sequence.

XX Sequence 327 BP; 63 A; 79 C; 116 G; 69 T; 0 U; 0 Other;

Query Match 100.0%; Score 327; DB 10; Length 327;
 Best Local Similarity 100.0%; Pred. No. 4.8e-80;
 Matches 327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGATCTGAATTCGAGTCTGCTTGGCTCGCGGTGAGGACAACTCTTCGGGTCTTT 60
 DB 1 AGATCTGAATTCGAGTCTGCTTGGCTCGCGGTGAGGACAACTCTTCGGGTCTTT 60

QY 61 CCACTACTCTTGGATCGGAACCCGTCGGCTCCGAACGGTACTCGGCACCGAGGGACC 120
 DB 61 CCACTACTCTTGGATCGGAACCCGTCGGCTCCGAACGGTACTCGGCACCGAGGGACC 120

QY 121 TGACGAGTCCGATCGACCGGATCGGAACCTCTTCGAGAAAGCGGTCTAACAGTCA 180
 DB 121 TGACGAGTCCGATCGACCGGATCGGAACCTCTTCGAGAAAGCGGTCTAACAGTCA 180

QY 181 AGTCGCAAGTATGGCTGAGCACCGTGGCGGCGGAGCGGGTGGCGGTGGTCTTC 240
 DB 181 AGTCGCAAGTATGGCTGAGCACCGTGGCGGCGGAGCGGGTGGCGGTGGTCTTC 240

QY 241 TGGCGGAGTGTGCTGATGATGATTAATTAAGTAGCGGTCTTGAGACGGCGGATGGT 300
 DB 241 TGGCGGAGTGTGCTGATGATGATTAATTAAGTAGCGGTCTTGAGACGGCGGATGGT 300

QY 301 AGGTGAGGTGTGGCGGTGAGATCT 327
 DB 301 AGGTGAGGTGTGGCGGTGAGATCT 327

RESULT 5
 AAV32372
 ID AAV32372 standard; DNA; 7469 BP.

XX AAV32372;

XX 25-MAR-2003 (revised)

XX 29-SEP-1998 (first entry)

XX Complete sequence of the pCLF plasmid.

KW Circular; adenovirus type 5; pCDNA3/Fiber plasmid; structural protein;
 KW complementation; fiber protein; gene therapy; HIV; tumour; early gene;
 KW Huntington's disease; Tay-Sachs disease; sickle cell disease;
 KW pCLF plasmid; AD2; adenovirus type 2; ds.

XX Synthetic.

XX Key Location/Qualifiers
 FT sig_peptide 907..1233
 FT /*tag= a
 FT /*note= "AD2 tripartite leader sequence"
 FT CDS 1237..2982
 FT /*tag= b
 FT /*product= "AD5 fiber protein"

XX W09813499-A2.

XX 02-APR-1998.

XX 24-SEP-1997; 97WO-EP005251.

XX 25-SEP-1996; 96US-00719806.

XX (NOVS) NOVARTIS AG.

XX (SCRI) SCRIPPS RES INST.

XX Nemerow GR, Von Seggern DJ;

XX WPI; 1998-230709/20.

XX Adenoviral vectors - which lack DNA encoding for structural protein or fibre protein used particularly for gene therapy.

XX Example 1; Page 85-94; 170pp; English.

XX The present sequence is that of a pCLF plasmid used in the method of the invention. The pCLF plasmid was derived from the pCDNA3/Fiber plasmid (AAV32371) containing an additional adenovirus type 2 (AD2) tripartite leader sequence to enhance expression. The pCLF plasmid also contains an adenovirus type 5 (AD5) fiber gene controlled by a CMV promoter and a neo resistant gene. The invention provides adenoviral vectors having deletions of all or part of various gene sequences encoding adenoviral structural proteins and/or early region proteins. Deletions in these proteins would allow a reduced risk of wild-type virus contamination and would also allow packaging of foreign DNA in such vectors for a variety of diagnostic and therapeutic applications. The adenoviral vectors having deletions in the structural and/or early gene regions are produced by cellular complementation of these adenoviral genes. Therefore, the pCLF plasmid was used as a complementation plasmid which was introduced into a host cell line where parts of the fiber gene region would be stably inserted into the host cell chromosomes. The resulting fiber gene deficient plasmid can be used as a gene delivery vector. The vectors can be used for diagnosis or gene therapy, e.g. for treating conditions characterised by hyper-proliferative cells (e.g. tumours), genetic diseases (e.g. Huntington's disease, Tay-Sachs disease, or sickle cell disease), or infections (e.g. HIV infection). They can also be used for in vitro production of biologically active proteins. (Updated on 25-MAR-2003 to correct PI field.)

XX Sequence 7469 BP; 1850 A; 1937 C; 1810 G; 1872 T; 0 U; 0 Other;

Query Match 99.4%; Score 325; DB 2; Length 7469;
 Best Local Similarity 100.0%; Pred. No. 3.7e-79;
 Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GATCTGAATTCGAGTCTGCTTGGCTCGCGGTGAGGACAACTCTTCGGGTCTTTC 61

DB 908 GATCTGAATTCGAGTCTGCTTGGCTCGCGGTGAGGACAACTCTTCGGGTCTTTC 967

QY 62 CAGTACTCTTGGATCGGAACCCGTCGGCTCTCGAACGGTACTTCGCCACCGAGGACCT 121

DB 968 CAGTACTCTTGGATCGGAACCCGTCGGCTCTCGAACGGTACTTCGCCACCGAGGACCT 1027

QY 122 GAGCGAGTCCGATCGACCCGATCGGAAACCTCTCGAGAAAGCGCTCTAACAGTCACA 181
DB 1028 GAGCGAGTCCGATCGACCCGATCGGAAACCTCTCGAGAAAGCGCTCTAACAGTCACA 1087
QY 182 GTCGCAAGGTAGCTGAGCACCGTGGCGGCGGCGGAGCGGTCGGGCTCTTCT 241
DB 1088 GTCGCAAGGTAGCTGAGCACCGTGGCGGCGGCGGAGCGGTCGGGCTCTTCT 1147
QY 242 GCGGAGGTGCTGCTGATGATGATTAATTAAGTAGGCGGTCTTGAGACGCGGATGGTGA 301
DB 1148 GCGGAGGTGCTGCTGATGATGATTAATTAAGTAGGCGGTCTTGAGACGCGGATGGTGA 1207
QY 302 GGTGAGGTGTCAGGCTTGAGATC 326
DB 1208 GGTGAGGTGTCAGGCTTGAGATC 1232

RESULT 6

AAAS9043
ID AAAS9043 standard; DNA; 7469 BP.

AC AAAS9043;

DT 15-SEP-2003 (revised)

DT 07-NOV-2000 (first entry)

DE Nucleotide sequence of a partial tripartite leader sequence.

KW Adenovirus; tripartite leader; adenovirus vector particle; gene delivery;
KW ss.

OS Human adenovirus type 5.

PN WO200042208-A1.

PD 20-JUL-2000.

PF 14-JAN-2000; 2000WO-EP000265.

PR 14-JAN-1999; 99US-0115920P.

XX (NOVS) NOVARTIS AG.

PA (NOVS) NOVARTIS-ERFINDUNGEN VERW GES MBH.

PA (SCRI) SCRIPPS RES INST.

PI Nemerow GR, Von Seggern DJ, Hallenbeck PL, Stevenson SC;

PI Skripchenko Y;

XX WPI; 2000-476068/41.

XX New nucleic acid comprising an adenovirus tripartite leader nucleotide
PT for producing high-capacity and targeted vectors for adenovirus-based
PT gene therapy.

PS Claim 10; Page 154-156; 212pp; English.

XX The specification describes a nucleic acid molecule comprising an
CC adenovirus (AV) tripartite leader (TPL) nucleotide with a sequence
CC comprising two different TPL exons or three same or different TPL exons.
CC The nucleic acid is used to produce an adenovirus vector particle,
CC deliver an exogenous gene to a target cell, pseudotype recombinant viral
CC vectors, target an adenovirus vector to a cell, produce a modified
CC adenovirus, deliver a heterologous gene to an animal and produce a
CC gutless adenoviral vector particle. The present sequence represents a
CC partial TPL sequence, which is used to construct nucleic acid molecules
CC of the invention. (Updated on 15-SEP-2003 to standardise OS field)

SQ Sequence 7469 BP; 1850 A; 1937 C; 1810 G; 1872 T; 0 U; 0 Other;

Query Match 99.4%; Score 325; DB 3; Length 7469;

Best Local Similarity 100.0%; Pred. No. 3.7e-79;

Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GATCTGAATTCGAGCTCGCTGTGGGTCGCGGTTGAGCAAACTCTTCGGGCTTTTC 61
DB 908 GATCTGAATTCGAGCTCGCTGTGGGTCGCGGTTGAGCAAACTCTTCGGGCTTTTC 967
QY 62 CAGTACTCTTCGATCGGAAACCCGCTCGGCTCCGACCGGTACTCCGCCACCGAGGACCT 121
DB 968 CAGTACTCTTCGATCGGAAACCCGCTCGGCTCCGACCGGTACTCCGCCACCGAGGACCT 1027
QY 122 GAGCGAGTCCGATCGACCCGATCGGAAACCTCTTCGAGAAAGCGCTCTAACAGTCACA 181
DB 1028 GAGCGAGTCCGATCGACCCGATCGGAAACCTCTTCGAGAAAGCGCTCTAACAGTCACA 1087
QY 182 GTCGCAAGGTAGCTGAGCACCGTGGCGGCGGCGGAGCGGTCGGGCTCTTCT 241
DB 1088 GTCGCAAGGTAGCTGAGCACCGTGGCGGCGGCGGAGCGGTCGGGCTCTTCT 1147
QY 242 GCGGAGGTGCTGCTGATGATGATTAATTAAGTAGGCGGTCTTGAGACGCGGATGGTGA 301
DB 1148 GCGGAGGTGCTGCTGATGATGATTAATTAAGTAGGCGGTCTTGAGACGCGGATGGTGA 1207
QY 302 GGTGAGGTGTCAGGCTTGAGATC 326
DB 1208 GGTGAGGTGTCAGGCTTGAGATC 1232

RESULT 7

ABA94252

XX ABA94252 standard; DNA; 7469 BP.

AC ABA94252;

DT 13-MAR-2002 (first entry)

DE Nucleotide sequence of expression plasmid pCLF.

XX Adenovirus; inverter terminal repeat sequence; ITRS; ocular disease;
KW fiber protein; photoreceptor; rhodopsin; stargardt disease gene; STDG1;
KW ophthalmological; antiinflammatory; antidiabetic; cytostatic;
KW gene therapy; ss.

OS Synthetic.

XX WO200183729-A2.

XX 08-NOV-2001.

XX 30-APR-2001; 2001WO-EP004863.

XX 01-MAY-2000; 2000US-00562934.

XX (NOVS) NOVARTIS AG.

PA (SCRI) SCRIPPS RES INST.

PA (NEME/) NEMEROW G R.

PA (VSEG/) VON SEGGERN D J.

PA (FRIE/) FRIEDLANDER M.

PI Nemerow GR, Von Seggern DJ, Friedlander M;

XX WPI; 2002-082846/11.

XX Polynucleotide for making vectors, useful for treating ocular diseases,
PT e.g., retinitis pigmentosa, comprises adenovirus inverter terminal repeat
PT sequences, packaging signal and photoreceptor-specific promoter.

PS Example 1; Page 108-110; 149pp; English.

XX The invention provides an isolated polynucleotide comprising adenovirus
CC (AV) inverter terminal repeat sequences (ITRS), AV packaging signal
CC operatively linked to ITRS and a photoreceptor-specific promoter. A
CC recombinant AV vector (AVV) comprising the polynucleotide is useful for
CC targeted delivery of a gene product to the eye (especially to the
CC vitreous cavity), for treating an ocular disease, e.g., retinal
CC degenerative disease, retinitis pigmentosa, Stargardt's disease, diabetic

CC retinopathies, retinal vascularizations, and retinoblastoma, of a mammal
 CC preferably human. The AAV comprises a fiber protein that specifically or
 CC selectively binds to receptors that are expressed on cells (preferably
 CC photoreceptors in the eye). Preferably, the recombinant virus comprise a
 CC fiber protein from an adenovirus type D subgroup or is a chimeric protein
 CC containing a portion of the N-terminus of an adenovirus type 2 or type 5
 CC penton, and the therapeutic product is a trophic factor, an anti-
 CC apoptotic factor, gene encoding a rhodopsin protein, a wild-type
 CC stargardt disease gene (STDG1), an anti-cancer agent and a protein that
 CC regulates expression of a photoreceptor specific gene product. The viral
 CC nucleic acid of AAV comprises ITRS and packaging signal derived from AAV
 CC subgroup B or C, especially an AV type 2 or type 5. AAV is also useful
 CC for targeted gene therapy, where the vector comprises an AV type 37 fiber
 CC protein or its portion, and selectively transduces photoreceptors and
 CC delivers a gene product encoded by AAV. The present sequence represents
 CC an expression plasmid pCLF containing the adenovirus 5 fiber gene
 XX
 SQ Sequence 7469 BP; 1850 A; 1937 C; 1810 G; 1872 T; 0 U; 0 Other;
 Query Match 99.4%; Score 325; DB 6; Length 7469;
 Best Local Similarity 100.0%; Pred. No. 3.7e-79;
 Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 GATCTGAATTCGAGCTCGCTGTGGCTCGCGTTTGAGGACAACTCTTCGGGTCCTTC 61
 Db 908 GATCTGAATTCGAGCTCGCTGTGGCTCGCGTTTGAGGACAACTCTTCGGGTCCTTC 967
 QY 62 CAGTACTCTTGGATCGGAACCCGTCGGCTCCGAACCGTACTCCGCCACCGAGGACCT 121
 Db 968 CAGTACTCTTGGATCGGAACCCGTCGGCTCCGAACCGTACTCCGCCACCGAGGACCT 1027
 QY 122 GAGCAGTCCGATCGACCGATCGGAACCTCTCGAAGAGGCGTCTAACAGTCA 181
 Db 1028 GAGCAGTCCGATCGACCGATCGGAACCTCTCGAAGAGGCGTCTAACAGTCA 1087
 QY 182 GTCCGAAGTAGCTGAGCACCCTGCGCGGGGAGCGGTGCGGTCGGGTTCTTCT 241
 Db 1088 GTCCGAAGTAGCTGAGCACCCTGCGCGGGGAGCGGTGCGGTCGGGTTCTTCT 1147
 QY 242 GCGGAGGTGCTGCTGATGATGTAATTAAGTAGCGGTCTTTGAGACGCGGATGTCGA 301
 Db 1148 GCGGAGGTGCTGCTGATGATGTAATTAAGTAGCGGTCTTTGAGACGCGGATGTCGA 1207
 QY 302 GGTGAGGTGTCGACGCTTGATC 326
 Db 1208 GGTGAGGTGTCGACGCTTGATC 1232
 RESULT 8
 ADB75149
 ID ADB75149 standard; DNA; 7469 BP.
 XX
 AC ADB75149;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Plasmid pCLF DNA sequence.
 XX
 KW ophthalmological; antiinflammatory; antidiabetic; gene therapy;
 KW adenovirus inverted terminal repeat sequence;
 KW adenovirus packaging signal; photoreceptor-specific promoter;
 KW adenovirus type 37; adenovirus type D serotype; adenovirus type 2;
 KW rhodopsin; wild-type Stargardt disease gene; STDG1; anti-apoptotic factor;
 KW retinal degenerative disease; retinitis pigmentosa; Stargardt's disease;
 KW diabetic retinopathy; retinal vascularisation; choroideraemia;
 KW gyrate atrophy; macular dystrophy; retinoblastoma;
 KW photoreceptor-restricted transgene expression;
 KW recombinant adenovirus vector; adenovirus type 5; fibre coding region;
 KW pCDNA3/Fibre; plasmid; cyclic; circular; ds;
 KW cucumber mosaic virus promoter; CMV promoter.
 XX
 OS Synthetic.

OS
 OS Human adenovirus type 5.
 XX Cucumber mosaic virus.
 PN
 XX US2002193327-A1.
 XX
 PD 19-DEC-2002.
 XX
 PF 01-MAY-2001; 2001US-00847101.
 XX
 PR 01-MAY-2000; 2000US-00562934.
 XX
 PA (SCRI) SCRIPPS RES INST.
 XX
 PI Nemerow GR, Von Seggern DJ, Friedlander M;
 XX WPI; 2003-657234/62.
 DR
 XX
 PT Novel nucleic acids comprising adenovirus inverted terminal repeat
 PT sequences, adenovirus packaging signals operatively linked to the
 PT retinitis pigmentosa.
 XX
 PS Example 1; Page 41-44; 106pp; English.
 XX
 CC The invention describes an isolated nucleic acid (I) comprising
 CC adenovirus inverted terminal repeat sequence, an adenovirus packaging
 CC signal operatively linked to the sequence, and a photoreceptor-specific
 CC promoter. A Recombinant adenovirus vector (II) comprising (I) is useful
 CC for targeted delivery of a gene product to the eye of a mammal which
 CC involves administering (II) that comprises heterologous DNA encoding the
 CC gene product or resulting in expression of the gene product, where the
 CC recombinant virus comprises a fibre protein that specifically or
 CC selectively binds to receptors that are expressed on cells which are
 CC photoreceptors, in the eye. The recombinant virus comprises a fibre
 CC protein which is an adenovirus type 37, from an adenovirus type D
 CC serotype. The fibre is a chimeric protein containing a sufficient portion
 CC of the N-terminus of an adenovirus type 2 or type 5 fibre protein for
 CC interaction with an adenovirus type 2 or type 5 penton, and a sufficient
 CC portion of an adenovirus serotype D knob portion of the fiber for
 CC selective binding to photoreceptors in the eye of a mammal. The
 CC encapsulated nucleic acid comprises a photoreceptor-specific promoter
 CC operatively linked to a nucleic acid comprising the therapeutic product
 CC which is chosen from trophic factor, anti-apoptotic factor, gene encoding
 CC a rhodopsin protein, wild-type Stargardt disease gene (STDG1), an anti-
 CC cancer agent and a protein that regulates expression of a photoreceptor-
 CC specific gene product. The delivery is effected for treatment of an
 CC ocular disease such as retinal degenerative disease e.g., retinitis
 CC pigmentosa, Stargardt's disease, diabetic retinopathies, retinal
 CC vascularisation, choroideraemia, gyrate atrophy or macular dystrophy or
 CC retinoblastoma inherited and acquired retinal and neovascular
 CC degenerative diseases. The viral nucleic acid comprises an adenovirus
 CC inverted terminal repeat (ITR) sequences, and an adenovirus packaging
 CC signal operatively linked to the sequence. The ITRs and packaging signal
 CC are derived from an adenovirus serotype B or C, or adenovirus type 2 or
 CC 5. The viral nucleic acid further comprises a photoreceptor-specific
 CC promoter. (II) includes photoreceptor promoters providing a means not
 CC only for specific targeting of expression in these cells, but also for
 CC photoreceptor-restricted transgene expression. This sequence represents a
 CC plasmid comprising an adenovirus fibre gene under the control of a
 CC cucumber mosaic virus (CMV) promoter.
 XX
 SQ Sequence 7469 BP; 1850 A; 1937 C; 1810 G; 1872 T; 0 U; 0 Other;
 Query Match 99.4%; Score 325; DB 10; Length 7469;
 Best Local Similarity 100.0%; Pred. No. 3.7e-79;
 Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 GATCTGAATTCGAGCTCGCTGTGGCTCGCGTTTGAGGACAACTCTTCGGGTCCTTC 61
 Db 908 GATCTGAATTCGAGCTCGCTGTGGCTCGCGTTTGAGGACAACTCTTCGGGTCCTTC 967
 QY 62 CAGTACTCTTGGATCGGAACCCGTCGGCTCCGAACCGTACTCCGCCACCGAGGACCT 121
 Db 968 GATCTGAATTCGAGCTCGCTGTGGCTCGCGTTTGAGGACAACTCTTCGGGTCCTTC 967
 QY 62 CAGTACTCTTGGATCGGAACCCGTCGGCTCCGAACCGTACTCCGCCACCGAGGACCT 121


```

FT      /*tag= a
FT      /note= "AD5 E4 regulatory gene"
FT      4051.. .4366
FT      /*tag= b
FT      /note= "AD5 leader sequence"
FT      4372.. .6124
FT      /*tag= c
FT      /note= "AD5 fiber gene"
XX
PN      WO9813499-A2.
XX
XX      02-APR-1998.
XX
XX      24-SEP-1997; 97WO-EP005251.
XX
XX      25-SEP-1996; 96US-00719806.
XX
XX      (NOVS ) NOVARTIS AG.
XX      (SCRI ) SCHIPPS RES INST.
XX
XX      Nemerow GR, Von Seggern DJ;
XX      WPI; 1998-230709/20.
XX
XX      Adenoviral vectors - which lack DNA encoding for structural protein or
XX      fibre protein used particularly for gene therapy.
XX
XX      Example 1; Page 131-145; 170pp; English.
XX
XX      The present sequence is that of a pE4/Fiber plasmid used in the method of
XX      the invention. The plasmid contains an adenovirus type 5 (AD5) fiber gene
XX      controlled by a CMV promoter, an AD5 E4 gene and an adenovirus type 2
XX      (AD2) tripartite leader sequence upstream of the fiber gene. The
XX      invention provides adenoviral vectors having deletions of all or part of
XX      various gene sequences encoding adenoviral structural proteins and/or
XX      early region proteins. Deletions in these proteins would allow a reduced
XX      risk of wild-type virus contamination and would also allow packaging of
XX      foreign DNA in such vectors for a variety of diagnostic and therapeutic
XX      applications. The adenoviral vectors having deletions in the structural
XX      and/or early gene regions are produced by cellular complementation of
XX      these adenoviral genes. Therefore, the pE4/Fiber plasmid was used as a
XX      complementation plasmid which was introduced into a host cell line where
XX      parts of the fiber and E4 gene region would be stably inserted into the
XX      host cell chromosomes. The resulting E4/fiber gene deficient plasmid can
XX      be used as a gene delivery vector. The vectors can be used for diagnosis
XX      or gene therapy, e.g. for treating conditions characterised by hyper-
XX      proliferative cells (e.g. tumours), genetic diseases (e.g. Huntington's
XX      disease, Tay-Sachs disease, or sickle cell disease), or infections (e.g.
XX      HIV infection). They can also be used for in vitro production of
XX      biologically active proteins. (Updated on 25-MAR-2003 to correct PI
XX      field.)
XX
XX      Sequence 10610 BP; 2807 A; 2821 C; 2446 G; 2536 T; 0 U; 0 Other;
XX
XX      Query Match      99.4%; Score 325; DB 2; Length 10610;
XX      Best Local Similarity 100.0%; Pred. No. 4e-79;
XX      Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX      2 GATCTGAATTCGAGTCGCTGTTGGGCTCGCGGTTGAGGACAAACTCTTCGGGTCCTTC 61
XX      4049 GATCTGAATTCGAGTCGCTGTTGGGCTCGCGGTTGAGGACAAACTCTTCGGGTCCTTC 4108
XX
XX      62 CAGTACTCTTGGATCGGAACCCGTCGGCTCCGAAACGGTACTCCGCCACCGAGGACCT 121
XX      4109 CAGTACTCTTGGATCGGAACCCGTCGGCTCCGAAACGGTACTCCGCCACCGAGGACCT 4168
XX
XX      122 GAGCGAGTCCGATCGACCGGATCGGAACCTCTCGAAGAGCGGTCTTAACCACTCACA 181
XX      4169 GAGCGAGTCCGATCGACCGGATCGGAACCTCTCGAAGAGCGGTCTTAACCACTCACA 4228
XX
XX      182 GTCCGAAGTAGGCTGAGCACCGTCGGCGGGCGGCGGCTCGCGGTTCTTCT 241
XX      4229 GTCCGAAGTAGGCTGAGCACCGTCGGCGGGCGGCGGCTCGCGGTTCTTCT 4288

```

```

QY      242 GCGGAGGTGCTGCTGATGATGTAATTAAAGTAGCGGCTCTTGAGACGGCGGATGCTCGA 301
DB      4289 GCGGAGGTGCTGCTGATGATGTAATTAAAGTAGCGGCTCTTGAGACGGCGGATGCTCGA 4348
QY      302 GGTGAGGTGTGGCAGGCTTGAGATC 326
DB      4349 GGTGAGGTGTGGCAGGCTTGAGATC 4373

RESULT 11
AAAS9051
ID      AAAS9051 standard; DNA; 10610 BP.
XX
XX      AC      AAAS9051;
XX
XX      DT      07-NOV-2000 (first entry)
XX
XX      DE      Nucleotide sequence of the E4/fiber-expressing plasmid pE4/Fiber.
XX
XX      KW      Adenovirus; tripartite leader; adenovirus vector particle; gene delivery;
XX      E4 gene; fiber gene; ss.
XX
XX      OS      Synthetic.
XX      Human adenovirus type 5.
XX
XX      PN      WO200042208-A1.
XX
XX      PD      20-JUL-2000.
XX
XX      PF      14-JAN-2000; 2000WO-EP000265.
XX
XX      PR      14-JAN-1999; 99US-0115920P.
XX
XX      PA      (NOVS ) NOVARTIS AG.
XX      (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.
XX      (SCRI ) SCHIPPS RES INST.
XX
XX      PI      Nemerow GR, Von Seggern DJ, Hallenbeck PL, Stevenson SC;
XX      Skripchenko Y;
XX      WPI; 2000-476068/41.
XX
XX      New nucleic acid comprising an adenovirus tripartite leader nucleotide
XX      for producing high-capacity and targeted vectors for adenovirus-based
XX      gene therapy.
XX
XX      Example 1; Page 164-167; 212pp; English.
XX
XX      The specification describes a nucleic acid molecule comprising an
XX      adenovirus (AV) tripartite leader (TPL) nucleotide with a sequence
XX      comprising two different TPL exons or three same or different TPL exons.
XX      The nucleic acid is used to produce an adenovirus vector particle,
XX      deliver an exogenous gene to a target cell, pseudotype recombinant viral
XX      vectors, target an adenovirus vector to a cell, produce a modified
XX      adenovirus, deliver a heterologous gene to an animal and produce a
XX      gutless adenoviral vector particle. The present sequence represents
XX      pE4/Fiber, a complementing plasmid containing E4 and fiber Adenoviral
XX      genes
XX
XX      Sequence 10610 BP; 2807 A; 2821 C; 2446 G; 2536 T; 0 U; 0 Other;
XX
XX      Query Match      99.4%; Score 325; DB 3; Length 10610;
XX      Best Local Similarity 100.0%; Pred. No. 4e-79;
XX      Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX      2 GATCTGAATTCGAGTCGCTGTTGGGCTCGCGGTTGAGGACAAACTCTTCGGGTCCTTC 61
XX      4049 GATCTGAATTCGAGTCGCTGTTGGGCTCGCGGTTGAGGACAAACTCTTCGGGTCCTTC 4108
XX
XX      62 CAGTACTCTTGGATCGGAACCCGTCGGCTCCGAAACGGTACTCCGCCACCGAGGACCT 121
XX      4109 CAGTACTCTTGGATCGGAACCCGTCGGCTCCGAAACGGTACTCCGCCACCGAGGACCT 4168

```


XX 01-MAY-2000; 2000US-00562934.
PR (SRI) SCHRIPS RES INST.
PA Nemerow GR, Von Seggern DJ, Friedlander M;
XX WPI; 2003-657234/62.
XX Novel nucleic acids comprising adenovirus inverted terminal repeat
PT sequences, adenovirus packaging signals operatively linked to the
PT sequences and photoreceptor-specific promoters, useful for treating
PT retinitis pigmentosa.
XX Example 1; Page 57-61; 106pp; English.
XX The invention describes an isolated nucleic acid (I) comprising
CC adenovirus inverted terminal repeat sequence, an adenovirus packaging
CC signal operatively linked to the sequence, and a photoreceptor-specific
CC promoter. A Recombinant adenovirus vector (II) comprising (I) is useful
CC for targeted delivery of a gene product to the eye of a mammal which
CC involves administering (II) that comprises heterologous DNA encoding the
CC gene product or resulting in expression of the gene product, where the
CC recombinant virus comprises a fibre protein that specifically or
CC selectively binds to receptors that are expressed on cells which are
CC photoreceptors, in the eye. The recombinant virus comprises a fibre
CC protein which is an adenovirus type 37, from an adenovirus type D
CC serotype. The fibre is a chimeric protein containing a sufficient portion
CC of the N-terminus of an adenovirus type 2 or type 5 fibre protein for
CC interaction with an adenovirus type 2 or type 5 penton, and a sufficient
CC portion of an adenovirus serotype D knob portion of the fiber for
CC selective binding to photoreceptors in the eye of a mammal. The
CC encapsulated nucleic acid comprises a photoreceptor-specific promoter
CC operatively linked to a nucleic acid comprising the therapeutic product
CC which is chosen from trophic factor, anti-apoptotic factor, gene encoding
CC a rhodopsin protein, wild-type Stargardt disease gene (STGD1), an anti-
CC cancer agent and a protein that regulates expression of a photoreceptor-
CC specific gene product. The delivery is effected for treatment of an
CC ocular disease such as retinal degenerative disease e.g., retinitis
CC pigmentosa, Stargardt's disease, diabetic retinopathies, retinal
CC vascularisation, choroidaemia, gyrate atrophy or macular dystrophy or
CC retinoblastoma inherited and acquired retinal and neovascular
CC degenerative diseases. The viral nucleic acid comprises an adenovirus
CC inverted terminal repeat (ITR) sequences, and an adenovirus packaging
CC signal operatively linked to the sequence. The ITRs and packaging signal
CC are derived from an adenovirus serotype B or C, or adenovirus type 2 or
CC 5. The viral nucleic acid further comprises a photoreceptor-specific
CC promoter. (II) includes photoreceptor promoters providing a means not
CC only for specific targeting of expression in these cells, but also for
CC photoreceptor-restricted transgene expression. This sequence represents a
CC plasmid expressing adenovirus type 5 fibre gene and E4 gene that can be
CC used to complement one or more delivery plasmids expressing E4 and fibre.
XX Sequence 10610 BP; 2807 A; 2821 C; 2446 G; 2536 T; 0 U; 0 Other;
SQ

Query Match 99.4%; Score 325; DB 10; Length 10610;
Best Local Similarity 100.0%; Pred. No. 4e-79;
Matches . 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GATCTGAATTCGAGTCGCTGTGGCTCGCGGTTGAGGACAACTCTTCGGGTCCTTC 61
DB 4049 GATCTGAATTCGAGTCGCTGTGGCTCGCGGTTGAGGACAACTCTTCGGGTCCTTC 4108
QY 62 CAGTACTCTTGGATCGGAACCGCTCGGCTCCGAAACGCTACTCCGCCACCGAGGACCT 121
DB 4109 CAGTACTCTTGGATCGGAACCGCTCGGCTCCGAAACGCTACTCCGCCACCGAGGACCT 4168
QY 122 GAGCGAGTCCGATCGACCGATCGGAACCTCTTCGAGAAAGGCTCTAACCACTACA 181
DB 4169 GAGCGAGTCCGATCGACCGATCGGAACCTCTTCGAGAAAGGCTCTAACCACTACA 4228
QY 182 GTCTCAAGTAGGCTGAGCACCCTGGCGCGGCGGAGCGGTCGGCGGTCGGGTCCTTTCT 241

Db 4229 GTCCCAAGTAGGCTGAGCACCGTGGCGGGCGGCGGAGCGGTCGGGTCGGGTTCT 4288
QY 242 GCGCGAGGTGCTGCTGATGATGTAATTAAAGTAGTGGCGGTCCTTGAGACGCGGATGGTCGA 301
Db 4289 GCGCGAGGTGCTGCTGATGATGTAATTAAAGTAGTGGCGGTCCTTGAGACGCGGATGGTCGA 4348
QY 302 GGTGAGCTGTGGCAGGCTTGAGATC 326
Db 4349 GGTGAGGTGTGGCAGGCTTGAGATC 4373
RESULT 14
ADF48802
ID ADF48802 standard; DNA; 10610 BP.
XX
AC ADF48802;
XX 12-FEB-2004 (first entry)
DE E4/fibre expressing plasmid pE1/fibre.
XX
KW cytostatic; anti-HIV; gene therapy; HIV gene expression inhibitor;
KW HIV gene expression activation; adenovirus tripartite leader; TPL;
KW gutless adenoviral vector particle;
KW helper-independent fiberless recombinant adenovirus vector;
KW packaging cell line; pseudotyping; adenovirus vector; gene therapy;
KW hereditary disorder; tumour; HIV infection; E4 transcription unit; fibre;
KW hygromycin resistance; ds; circular; cyclic.
OS Synthetic.
OS Human adenovirus type 5.
FN US2003157688-A1.
XX 21-AUG-2003.
XX 14-JAN-2000; 2000US-00482682.
PR 14-JAN-1999; 99US-0115920P.
PR 26-JUN-2000; 2000US-00423783.
PA (VSEG/) VON SEGGERN D J.
PA (NEME/) NEMEROW G R.
PA (HALL/) HALLENBECK P.
PA (STEV/) STEVENSON S.
PA (SKRI/) SKRIPCHENKO Y.
PI Von Seggern DJ, Nemerow GR, Hallenbeck P, Stevenson S;
XX Skripchenko Y;
DR WPI; 2003-843463/78.
XX Novel isolated nucleic acid molecule useful for delivering heterologous
PT gene to human or any animal, or for producing gutless adenoviral vector
PT particle.
XX Example 1; SEQ ID NO 16; 157pp; English.
XX The invention describes an isolated nucleic acid molecule (I) comprising
CC an adenovirus tripartite leader (TPL) nucleotide, the TPL nucleotide
CC sequence comprising a first and second different TPL exons or first
CC second and third same or different TPL exons, the TPL exons chosen from
CC complete or partial TPL exon 1, complete TPL exon 2 and complete TPL exon
CC 3. (I) is useful for delivering a heterologous gene to a human or any
CC animal, or for producing a gutless adenoviral vector particle. A
CC recombinant adenovirus particle (II) is useful for delivery of an
CC exogenous gene to a target cell which involves contacting the cell with
CC an amount of (I) sufficient to infect the cell. A helper-independent
CC fiberless recombinant adenovirus vector genome (III) is useful for
CC producing an adenovirus vector particle containing (III) which involves
CC providing a packaging cell line which complements replication and
CC packaging of the genome and (III) which is deficient in expressing
CC sufficient functional fiber protein to support assembly of fiber

CC containing particles and harvesting the particle produced by the cell
CC line. (III) is useful for pseudotyping recombinant viral vectors which
CC involves complementing a missing fiber gene of (III) or helper-dependent
CC fiberless recombinant adenovirus vector genome by expressing in packaging
CC cells a fiber gene from a different adenoviral serotype than the
CC recombinant adenovirus vector. (III) is also useful for specifically
CC targeting an adenovirus vector to a cell of choice. (I) is useful for
CC gene therapy. (II) is useful for treating diseases such as hereditary
CC disorder, and for reducing proliferation of tumour cells in a subject, or
CC to disrupt HIV infection. This sequence represents the complementing
CC plasmid pE4/fibre that expresses the adenoviral E4 and fibre genes.
XX
SQ Sequence 10610 BP; 2807 A; 2821 C; 2446 G; 2536 T; 0 U; 0 Other;

Query Match 99.4%; Score 325; DB 10; Length 10610;
Best Local Similarity 100.0%; Pred. No. 4e-79;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 GATCTGAATTCGAGTCTGCTGTTGGCTCGCGGTTGAGACAAACTCTTCGGGCTTTTC 61
Db 4049 GATCTGAATTCGAGTCTGCTGTTGGCTCGCGGTTGAGACAAACTCTTCGGGCTTTTC 4108
QY 62 CAGTACTCTTGGATCGGAACCGCTCGGCTCGGACGGTACTCCGCCACCGAGGACCT 121
Db 4109 CAGTACTCTTGGATCGGAACCGCTCGGCTCGGACGGTACTCCGCCACCGAGGACCT 4168
QY 122 GAGCGAGTCCGCATCGACCGATCGGAACCTCTCGAGAAAGGCGTCTAACCACTCACA 181
Db 4169 GAGCGAGTCCGCATCGACCGATCGGAACCTCTCGAGAAAGGCGTCTAACCACTCACA 4228
QY 182 GTCCGAAGTAGGCTGAGCACCCTGCGGCGCGGACGCGGTGCGGCTGGGGTTTTC 241
Db 4229 GTCCGAAGTAGGCTGAGCACCCTGCGGCGCGGACGCGGTGCGGCTGGGGTTTTC 4288
QY 242 GCGGAGGTGCTGCTGATCATTAATAAGTAGGCGGTCTTGACACCGCGATGTCGA 301
Db 4289 GCGGAGGTGCTGCTGATCATTAATAAGTAGGCGGTCTTGACACCGCGATGTCGA 4348
QY 302 GGTGAGGTGTGGCAGGCTTGAGATC 326
Db 4349 GGTGAGGTGTGGCAGGCTTGAGATC 4373

RESULT 15
AAV32374/c
ID AAV22374 standard; DNA; 14455 BP.

XX AAV32374;

XX 25-MAR-2003 (revised)

DT 29-SEP-1998 (first entry)

XX Complete sequence of the pE1/Fiber plasmid.

KW Circular; adenovirus type 5; pE1/Fiber plasmid; structural protein;
KW complementation; fiber protein; gene therapy; HIV; tumour; AD5;
KW early gene; Huntington's disease; Tay-Sachs disease; sickle cell disease;
KW E1 regulatory protein; ds.

OS Synthetic.

XX Key Location/Qualifiers
FH 1460. .4998
FT misc_feature /tag= a
FT /note= "AD5 E1 regulatory gene"
FT complement (10922. .14223)
FT /tag= b
FT /note= "AD5 fiber gene consisting of a CMV promoter at 5'
FT end of this gene"

XX WO9813499-A2.

XX 02-APR-1998.

XX 24-SEP-1997; 97WO-EP005251.
XX 25-SEP-1996; 96US-00719806.
XX (NOVS) NOVARTIS AG.
XX (SCKI) SCRIPPS RES INST.
XX Nemerow GR, Von Seggern DJ;
XX WPI, 1998-230709/20.
XX
XX Adenoviral vectors - which lack DNA encoding for structural protein or
XX fibre protein used particularly for gene therapy.
XX
XX Example 1; Page 112-131; 170pp; English.

CC The present sequence is that of a pE1/Fiber plasmid used in the method of
CC the invention. The plasmid contains an adenovirus type 5 (AD5) fiber gene
CC controlled by a CMV promoter, an AD5 E1 gene and a pMAM backbone. The
CC invention provides adenoviral vectors having deletions of all or part of
CC various gene sequences encoding adenoviral structural proteins and/or
CC early region proteins. Deletions in these proteins would allow a reduced
CC risk of wild-type virus contamination and would also allow packaging of
CC foreign DNA in such vectors for a variety of diagnostic and therapeutic
CC applications. The adenoviral vectors having deletions in the structural
CC and/or early gene regions are produced by cellular complementation of
CC these adenoviral genes. Therefore, the pE1/Fiber plasmid was used as a
CC complementation plasmid which was introduced into a host cell line where
CC parts of the fiber and E1 gene region would be stably inserted into the
CC host cell chromosomes. The resulting E1/fiber gene deficient plasmid can
CC be used as a gene delivery vector. The vectors can be used for diagnosis
CC or gene therapy, e.g. for treating conditions characterised by hyper-
CC proliferative cells (e.g. tumours), genetic diseases (e.g. Huntington's
CC disease, Tay-Sachs disease, or sickle cell disease), or infections (e.g.
CC HIV infection). They can also be used for in vitro production of
CC biologically active proteins. (Updated on 25-MAR-2003 to correct PI
CC field.)

SQ Sequence 14455 BP; 3698 A; 3271 C; 3565 G; 3921 T; 0 U; 0 Other;

Query Match 99.4%; Score 325; DB 2; Length 14455;
Best Local Similarity 100.0%; Pred. No. 4.4e-79;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GATCTGAATTCGAGTCTGCTGTTGGCTCGCGGTTGAGACAAACTCTTCGGGCTTTTC 61
Db 13315 GATCTGAATTCGAGTCTGCTGTTGGCTCGCGGTTGAGACAAACTCTTCGGGCTTTTC 13256
QY 62 CAGTACTCTTGGATCGGAACCGCTCGGCTCTCGAACCGTACTCCGCCACCGAGGACCT 121
Db 13255 CAGTACTCTTGGATCGGAACCGCTCGGCTCTCGAACCGTACTCCGCCACCGAGGACCT 13196
QY 122 GAGCGAGTCCGCATCGACCGATCGGAACCTCTCGAGAAAGGCGTCTAACCACTCACA 181
Db 13195 GAGCGAGTCCGCATCGACCGATCGGAACCTCTCGAGAAAGGCGTCTAACCACTCACA 13136
QY 182 GTCCGAAGTAGGCTGAGCACCCTGCGGCGCGGACGCGGTGCGGCTGGGGTTTTC 241
Db 13135 GTCCGAAGTAGGCTGAGCACCCTGCGGCGCGGACGCGGTGCGGCTGGGGTTTTC 13076
QY 242 GCGGAGGTGCTGCTGATGATTAATAAGTAGGCGGTCTTGAGACCGCGGATGTCGA 301
Db 13075 GCGGAGGTGCTGCTGATGATTAATAAGTAGGCGGTCTTGAGACCGCGGATGTCGA 13016
QY 302 GGTGAGGTGTGCGAGGCTTGAGATC 326
Db 13015 GGTGAGGTGTGCGAGGCTTGAGATC 12991

Search completed: July 14, 2005, 07:01:34
Job time : 465.082 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 05:15:57 ; Search time 3113.52 Seconds
(without alignments)
3997.736 Million cell updates/sec

Title: US-09-482-682-26

Perfect score: 327

Sequence: 1 agatctgaattcgagctgcg.....gtgtgcaggcttgagatct 327

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hic:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gsl1:*

9: gb_gsl2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	43.2	13.2	1086	9	CL467818 SAIL 1276
C 2	42.6	13.0	1262	9	CL496190 SAIL_620
C 3	41.6	12.7	1164	8	CC258768 CH261-164
C 4	41.6	12.7	1164	8	CC258769 CH261-164
C 5	40.4	12.4	710	4	BF972939 602241236
C 6	40.2	12.3	364	1	AJ602111 AJ602111
C 7	40.2	12.3	632	6	CA731965 wlpic.pk0
C 8	40.2	12.3	598	6	CA378248 657011 NC
C 9	40	12.2	1177	8	CC286173 CH261-29F
C 10	39.6	12.1	936	9	CL513605 SAIL 877
C 11	39.2	12.0	511	1	AJ437840 AJ437840
C 12	39.2	12.0	629	6	CA728476 wdiic.pk0
C 13	39.2	12.0	719	8	BZ715587 OGEAT34TC
C 14	39.2	12.0	847	5	BX900296 BX900296
C 15	39	11.9	378	6	CB685324 OSUNEf15K
C 16	39	11.9	608	6	CB648908 OSUNEb121
C 17	39	11.9	652	6	CB665091 OSUNEb11D
C 18	39	11.9	703	6	CB662732 OSUNEb07A
C 19	39	11.9	726	9	AG101098 Pan trogl
C 20	39	11.9	831	7	CF711474 CCAD883TR
C 21	39	11.9	885	9	CC697699 OGNAJ58TH
C 22	39	11.9	1224	9	CL974456 OseIFCC025
C 23	38.6	11.8	384	1	AJ602280 AJ602280
C 24	38.6	11.8	621	6	CA733511 wlpic.pk0

C 25	38.6	11.8	622	4	BJ280072
C 26	38.6	11.8	649	6	CD875521
C 27	38.6	11.8	672	6	CD884257 F1.116A21
C 28	38.6	11.8	701	6	CD875244 AZ03.104K
C 29	38.2	11.7	909	9	AG125251 Pan trogl
C 30	38.2	11.7	927	9	CL467547 SAIL_1271
C 31	38.2	11.7	1276	9	CL491442 SAIL_555
C 32	38	11.6	303	9	CL256246 FHCRC-GT-
C 33	38	11.6	684	6	CD464049 EL01T0206
C 34	38	11.6	896	6	CD439437 EL01N0524
C 35	38	11.6	1978	3	CNS0A7EC Arabidops
C 36	37.8	11.6	769	3	CNS032YU Tetraodon
C 37	37.8	11.6	913	9	CL473483 SAIL_201
C 38	37.8	11.6	943	9	CL466335 SAIL_1254
C 39	37.8	11.6	1058	4	BG387202 602455932
C 40	37.6	11.5	1063	5	BX331486 BX331486
C 41	37.6	11.5	397	1	AJ437938 AJ437938
C 42	37.6	11.5	487	2	BE230675 99AS897 R
C 43	37.6	11.5	568	6	CD205709 HSI_15_G0
C 44	37.6	11.5	661	2	BE823637 GM700021A
C 45	37.6	11.5	724	6	CB627516 OS11EB021

ALIGNMENTS

RESULT 1	CL467818/c	CL467818	1086 bp	DNA	linear	GSS 31-MAR-2004
LOCUS	SAIL_1276_G09.v1	SAIL_1276_G09.v1	SAIL Collection	Arabidopsis thaliana	genomic clone	
DEFINITION	SAIL_1276_G09.v1	SAIL_1276_G09.v1	genomic survey sequence.			
ACCESSION	CL467818	CL467818	GI:45870723			
VERSION	GSS.					
KEYWORDS	Arabidopsis thaliana (thale cress)					
SOURCE	Arabidopsis thaliana					
ORGANISM	Arabidopsis thaliana					
REFERENCE	1 (bases 1 to 1086)					
AUTHORS	Sessions,A., Burke,E., Presting,G., Aux,G., McElver,J., Patton,D., Dietrich,B., Ho,P., Bacwaden,J., Ko,C., Clarke,J.D., Cotton,D., Bullis,D., Snell,J., Miguel,T., Hutchison,D., Kimmerly,B., Mitzel,T., Katagiri,F., Glazebrook,J., Law,M. and Goff,S.A.					
TITLE	A high-throughput Arabidopsis reverse genetics system					
JOURNAL	Plant Cell 14 (12), 2985-2994 (2002)					
MEDLINE	22356987					
PUBMED	12468722					
COMMENT	Contact: Sessions A Applied Trait Genetics Syngenta Biotechnology Inc. 3054 Cornwallis Rd., Research Triangle Park, NC 27709, USA Email: allen.sessions@syngenta.com ABRC Stock Number CS847449; T-DNA left border flanking sequences of Syngenta Arabidopsis Insertion Library (SAIL) lines are available through the Arabidopsis Biological Resource Center (ABRC). Sequences represent a pool of amplified genomic regions and not single contiguous sequences. Class: TDNA tagged. Location/Qualifiers 1. .1086 /organism="Arabidopsis thaliana" /mol_type="genomic DNA" /ecotype="Columbia" /db xref="taxon:3702" /clone="SAIL_1276_G09.v1" /clone.lib="SAIL Collection" /note="T-DNA left border sequences were isolated using a modified TAIL-PCR strategy"					

FEATURES

ORIGIN

Query Match 13.2%; Score 43.2; DB 9; Length 1086;
Best Local Similarity 56.1%; Pred. No. 0.41;

Matches 78; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 179 ACAGTCGAAGTAGCTGAGCACCGTGGCGGCGGACCGGTGCGGTCTGGGTCTTT 238
 |||||
 Db 570 AGAGTGGGAAGGTGCGAGGAGTAGAGTGTGGAAGGAGGAGGCGGCGGCGGCGGCGT 511
 |||||

QY 239 TCTGCGGAGGCTGCTGATGATGTAATTAAGTAGGCGGTCTTGAGACGCGGATGGT 298
 |||||
 Db 510 GGGGAGGAGGAGAGAGGGGGGNGGGGGGGGGGGGGGGGGGGGGGGGGGGGGTGA 451
 |||||

QY 299 CGAGGTGAGGTGTGGCAGG 317
 |||||

Db 450 GGAGGGGAGGGGGGAGG 432
 |||||

RESULT 2
 CL496190/c
 LOCUS
 DEFINITION SAIL_620_G10.v3 SAIL Collection Arabidopsis thaliana genomic clone
 SAIL_620_G10.v3, genomic survey sequence.

ACCESSION
 VERSION
 CL496190.1 GI:45988256

KEYWORDS
 GSS.

SOURCE
 ORGANISM Arabidopsis thaliana (thale cress)

Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE
 1 (bases 1 to 1262)
 AUTHORS Sessions, A., Burke, E., Presting, G., Aux, G., McElver, J., Patton, D.,
 Dietrich, B., Ho, P., Bacwaden, J., Ko, C., Clarke, J. D., Cotton, D.,
 Bullis, D., Snell, J., Miguel, T., Hutchison, D., Kimmerly, B.,
 Mitzel, T., Katagiri, F., Glazebrook, J., Law, M. and Goff, S. A.
 A high-throughput Arabidopsis reverse genetics system
 Plant Cell 14 (12), 2985-2994 (2002)

TITLE
 MEDLINE
 PUBMED

COMMENT

Contact: Sessions A
 Applied Trait Genetics
 Syngenta Biotechnology Inc.
 3054 Cornwallis Rd., Research Triangle Park, NC 27709, USA
 Email: allen.sessions@syngenta.com
 ABRC Stock Number C5826573; T-DNA left border flanking sequences of
 Syngenta Arabidopsis Insertion Library (SAIL) lines are available
 through the Arabidopsis Biological Resource Center (ABRC).
 Sequences represent a pool of amplified genomic regions and not
 single contiguous sequences.
 Class: TDNA tagged.

FEATURES
 source

1..1262
 Location/Qualifiers
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /ecotype="Columbia"
 /db_xref="taxon:3702"
 /clone="SAIL_620_G10.v3"
 /clone_lib="SAIL Collection"
 /note="T-DNA left border sequences were isolated using a
 modified TAIL-PCR strategy"

ORIGIN

Query Match 13.0%; Score 42.6; DB 9; Length 1262;
 Best Local Similarity 57.7%; Pred. No. 0.61;
 Matches 75; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 189 GGTAGCTGAGCACCGTGGCGGCGGCGGCGGTGCGGTGTTTCTGCGCGAG 248
 |||||

Db 1217 GGGGGCGGGGGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGTTCGCGCGC 1159
 |||||

QY 249 GTGCTGCTGATGATGTAATTAAGTAGGCGGTCTTGAGACGCGCGATGTCGAGGTGAGG 308
 |||||

Db 1157 GTNTTGGTGGGGCGGTGCGGTGGGTGGCGGTGGGGCGCGCGGGGGGGGGGGG 1098
 |||||

QY 309 TGTGGCAGCG 318

Db 1097 TGGGCGCGGC 1088
 |||||

RESULT 3

CC258768/c

LOCUS

DEFINITION CC258768 1164 bp DNA linear GSS 13-MAY-2003
 CH261-164A7_RM1.1 CH261 Gallus gallus genomic clone CH261-164A7,
 genomic survey sequence.

ACCESSION
 VERSION
 CC258768

KEYWORDS
 GSS.

SOURCE
 ORGANISM Gallus gallus (chicken)

Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.

REFERENCE
 1 (bases 1 to 1164)

AUTHORS Kremitzki, C., Higginbotham, J., Wylie, K., Carter, J., McPherson, J.,
 Warren, W., Graves, T., Mardis, E. and Wilson, R.

TITLE
 JOURNAL
 COMMENT

Unpublished (2003)

Contact: Richard K. Wilson

Genome Sequencing Center

Washington University School of Medicine

Email: submissions@wustl.edu

Insert Length: 182000 Std Error: 0.00

Seq primer: RM1 TAGCACTCACTATAGGAGGA

Class: BAC ends

High quality sequence start: 186

High quality sequence stop: 475.

Location/Qualifiers

1..1164

/organism="Gallus gallus"

/mol_type="genomic DNA"

/strain="Red Jungle Fowl"

/db_xref="taxon:9031"

/clone="CH261-164A7"

/sex="female"

/cell_line="UCD001, inbred 256"

/clone_lib="CH261"

/note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
 CH261 Female Chicken library - for library and clone
 ordering information: http://www.chori.org/bacpac"

ORIGIN

Query Match 12.7%; Score 41.6; DB 8; Length 1164;

Best Local Similarity 59.2%; Pred. No. 1.1;

Matches 71; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 204 GTGCGGCGCGACCGGTGCGGTGCGGTGTTTCTGCGGAGGTGCTGCTGATGATG 263
 |||||

Db 741 GCGTGTGGGTGGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTG 682
 |||||

QY 264 TAATTAAGTAGGCGGTCTTGAGACGCGGATGTCGAGGTGAGGTGCGCAGGCTTCAG 323
 |||||

Db 681 GGGTGTGGGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 622
 |||||

RESULT 4

CC258769/c

LOCUS

DEFINITION CC258769 1164 bp DNA linear GSS 13-MAY-2003
 CH261-164A7_RM1.2 CH261 Gallus gallus genomic clone CH261-164A7,
 genomic survey sequence.

ACCESSION
 VERSION
 CC258769

KEYWORDS
 GSS.

SOURCE
 ORGANISM Gallus gallus (chicken)

Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.

REFERENCE
 1 (bases 1 to 1164)

Db 247 GGGCTTGACCACTCGACAGAGCGGGTCGGAGCGGTCTTGGGGAAGGCGACGCTTCG 188
 QY 176 GTACAGTCGCAAGTAGCTGAGCACCGTGGCGGGCGGACGCGGTGGCGGTTCGGGGTT 235
 Db 187 GACCTGGTGGCGGGCGGTGATCGCTGGTGTAGAGCATGACAGGGCTGCGTGGGGAG 128
 QY 236 GTTCTTGGCGGAG 248
 Db 127 GATGGTGGCGGAG 115

RESULT 7

CA731965/c
 LOCUS CA731965 632 bp mRNA linear EST 26-NOV-2002
 DEFINITION wlpic.pk002.j18 wlpic Triticum aestivum cDNA clone wlpic.pk002.j18
 5' end, mRNA sequence.

ACCESSION CA731965
 VERSION CA731965
 KEYWORDS CA731965.1 GI:25547563
 ORGANISM Triticum aestivum (bread wheat)
 SOURCE Triticum aestivum
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Poideae; Triticeae; Triticum.

REFERENCE 1 (bases 1 to 632)
 AUTHORS Tingey,S.V., Powell,W., Wolters,P., Dolan,M., Hainey,C., Yuan,Z.,
 Miao,G., Caraher,N. and Hanafey,M.K.
 TITLE DuPont Wheat cDNA Sequence
 JOURNAL Unpublished (2002)
 COMMENT Contact: Scott V. Tingey
 Crop Genetics

1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
 Tel: 302-631-2602
 Fax: 302-631-2607
 Email: Scott.V.Tingey@USA.dupont.com
 Seq primer: M13.

FEATURES

source
 1..632
 /location/Qualifiers
 /organism="Triticum aestivum"
 /mol_type="mRNA"
 /db_xref="taxon:4565"
 /clone="wlpic.pk002.j18"
 /tissue_type="lemma and palea"
 /lab_host="DH10B"
 /clone_lib="wlpic"
 /note="Vector: pBluescript SK+; Site_1: EcoRI; Site_2:
 XhoI; Wheat (Triticum aestivum, Hi Line) lemma and palea"

ORIGIN

Query Match 12.3%; Score 40.2; DB 6; Length 632;
 Best Local Similarity 56.4%; Pred. No. 2.6;
 Matches 75; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
 QY 116 GGACCTGAGCGAGTCGGATCGACCGGATCGGAAACCTCTCGAGAAAGGGTCTTAACA 175
 Db 546 GGGCTTGACCACTCGACAGAGCGGGTCGGAGGGTCTTGGGGAAGGCGACGCTTCG 487
 QY 176 GTACAGTCGCAAGTAGCTGAGCACCGTGGCGGGCGGACGCGGTGGCGGTTCGGGGTT 235
 Db 486 GACCTGGTGGCGGGCGGTGATCGCTGGTGTAGAGCATGACAGGGCTGCGGTGCGGGAG 427
 QY 236 GTTCTTGGCGGAG 248
 Db 426 GATGGTGGCGGAG 414

RESULT 8

CA378248/c
 LOCUS CA378248 598 bp mRNA linear EST 06-NOV-2002
 DEFINITION 657011 NCCOWA 1RT Oncorhynchus mykiss cDNA clone 1RT42L01_B_F01 5',
 mRNA sequence.

ACCESSION CA378248
 VERSION CA378248.1
 KEYWORDS GI:24697932
 SOURCE EST.
 ORGANISM Oncorhynchus mykiss (rainbow trout)

Oncorhynchus mykiss
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 1 (bases 1 to 598)
 REFERENCE 1 (bases 1 to 598)
 AUTHORS Rexroad,C.E. 3rd, Lee,Y., Keele,J.W., Karamycheva,S., Brown,G.,
 Koop,B., Gahr,S.A., Palti,Y. and Quackenbush,J.
 TITLE Sequence analysis of a rainbow trout cDNA library and creation of a
 gene index

JOURNAL Cytogenet. Genome Res. 102 (1-4), 347-354 (2003)
 COMMENT CONTACT: Rexroad CE
 USDA, ARS, National Center for Cool and Cold Water Aquaculture
 11876 Lestown Road, Kearneysville, WV 25430, USA
 Tel: 304 724 8340 x2129
 Fax: 304 725 0351
 Email: crexroad@nccowa.ars.usda.gov
 Single pass sequencing. Bases called with phred v0.020425.c and
 trimmed with the aid of the trim_alt option. Vector identified by
 cross match v0.990329.

Seq primer: AGCGGATAACAATTTCACACAGGA.

FEATURES

source
 1..598
 /location/Qualifiers
 /organism="Oncorhynchus mykiss"
 /mol_type="mRNA"
 /db_xref="taxon:8022"
 /clone="1RT42L01_B_F01"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /clone_lib="NCCOWA 1RT"
 /note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
 Library made from pooled tissue from brain, gill, liver,
 spleen, muscle, and kidney."

ORIGIN

Query Match 12.2%; Score 40; DB 6; Length 598;
 Best Local Similarity 50.0%; Pred. No. 3;
 Matches 100; Conservative 0; Mismatches 100; Indels 0; Gaps 0;
 QY 111 CCGAGGACCTGACGAGTCGGATCGACCGATCGGAAACCTCTCGAGAAAGCGCTCT 170
 Db 308 CGGTAGAATATATGATCGTGGAACTCTGGGGGCGAGCTCTCAAGCGGTGATGACGCTACG 249
 QY 171 AACCACTCACAGTCGCAAGGTAGGTGAGCACCGCTGGCGGGCGGCGGCTGGCGGTGCG 230
 Db 248 GACCGGTTTCCTCGAAGGCTTCCCGAGTGGGTGTCGATGTCAGCTTGACCTGGTTCG 189
 QY 231 GGGTTGTTTCTGGCGGAGGTGCTCTGATGATGTAATTAAGTAGGGCGGTCTTGAGACGG 290
 Db 188 CTGTCTTGGCGAGCTTGTGTCGCCCTCGATGTCGATGTAAGGAGGCGCTTGAGGCGCAG 129
 QY 291 CGGATGTCGAGGTGAGGTG 310
 Db 128 GAGAAGCGGAGGCCAGATG 109

RESULT 9

CC286173/c
 LOCUS CC286173 1177 bp DNA linear GSS 13-MAY-2003
 DEFINITION CH261-29F4_RM1.1 CH261 Gallus gallus genomic clone CH261-29F4,
 genomic survey sequence.

ACCESSION CC286173
 VERSION CC286173.1
 KEYWORDS GSS.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 1 (bases 1 to 1177)

AUTHORS
 Krenitzki,C., Higginbotham,J., Wylie,K., Carter,J., McPherson,J.,
 Warren,W., Graves,T., Mardis,E. and Wilson,R.
TITLE
 Gallus gallus BAC End Reads
JOURNAL
 Unpublished (2003)
COMMENT
 Contact: Richard K. Wilson
 Genome Sequencing Center
 Washington University School of Medicine
 Email: submissions@watson.wustl.edu
 Insert Length: 182000 Std Error: 0.00
 Seq primer: RM1 TACCACTCACTATAGGGAGA
 Class: BAC ends
 High quality sequence start: 135
 High quality sequence stop: 281.

FEATURES
 source
 1..1177
 Location/Qualifiers
 /organism="Gallus gallus"
 /mol_type="genomic DNA"
 /strain="Red Jungle Fowl"
 /db_xref="taxon:9031"
 /clone="CH261-29F4"
 /sex="female"
 /cell_line="UCD001, inbred 256"
 /clone_lib="CH261"
 /note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
 CH261 Female Chicken library - for library and clone
 ordering information: http://www.chori.org/bacpac"

ORIGIN
 Query Match 12.2%; Score 40; DB 8; Length 1177;
 Best Local Similarity 55.9%; Pred. No. 3.2;
 Matches 76; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 189 GTAGGCTGACACCGTGGCGGGCGAGCGGTGGCGGTGGTGTCTGGCGGAG 248
 |||||
 DB 725 GGGAGGGAGGGGGGTGGAGGGGGGAGGGGTGGAGGTGGGGGGGAGAGGGGAG 666
 |||||
 QY 249 GTGCTGCTGATGATTAATTAAGTAGGCGGTCTTGAGACGCGGATGTCGAGGTGAGG 308
 |||||
 DB 665 GGGAGGGGTGGGATGGGGGTAAATGGGGGGGCTAGAGAGGGCGAGGGTGGGGGAGAG 606
 |||||
 QY 309 TGTGCAGGCTTGACA 324
 |||||
 DB 605 GGGGGGGGGGGGGGA 590
 |||||

RESULT 10
 CL513605/c
 LOCUS
 DEFINITION
 SAIL_877_H06.v1 SAIL Collection Arabidopsis thaliana genomic clone
 CL513605_1 GI:46010925
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Arabidopsis thaliana (thale cress)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
 1 (bases 1 to 936)
 Sessions,A., Burke,E., Presting,G., Aux,G., McElver,J., Patton,D.,
 Dietrich,B., Ho,P., Bacwaden,J., Ko,C., Clarke,J.D., Cotton,D.,
 Bullis,D., Snell,J., Miguel,T., Hutchison,D., Kimmerly,B.,
 Mitzel,T., Katagiri,F., Glazebrook,J., Law,M. and Goff,S.A.
 A high-throughput Arabidopsis reverse genetics system
 Plant Cell 14 (12), 2985-2994 (2002)
 22356987
 12468722
 Contact: Sessions A
 Applied Trait Genetics
 Syngenta Biotechnology Inc.
 3054 Cornwallis Rd., Research Triangle Park, NC 27709, USA
 Email: allen.sessions@syngenta.com
 ABRC Stock Number CS839484; T-DNA left border flanking sequences of

FEATURES
 source
 1..936
 Location/Qualifiers
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /ecotype="Columbia"
 /db_xref="taxon:3702"
 /clone="SAIL_877_H06.v1"
 /clone_lib="SAIL Collection"
 /note="T-DNA left border sequences were isolated using a
 modified TAIL-PCR strategy"

ORIGIN
 Query Match 12.1%; Score 39.6; DB 9; Length 936;
 Best Local Similarity 56.0%; Pred. No. 4;
 Matches 75; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 188 AGTAGGCTGAGCACCGTGGCGGGCGGCGAGCGGTGGCGGTGGTGTCTTCTGGCGGA 247
 |||||
 DB 875 AGATGGGGGAGCCCGAGGGGGGGGGGAGCGGGGGGGGAGTGGGGGGGGGGGG 816
 |||||
 QY 248 GGTGCTGCTGATGATTAATTAAGTAGGCGGTCTTGAGACGCGGATGTCGAGGTGAG 307
 |||||
 DB 815 GGGGAGATGCTGTAGGCAAGAGATGGGACACAGGGGAGGTGGGTAGGGGGGGGG 756
 |||||
 QY 308 GTGTGGCAGGCTTG 321
 |||||
 DB 755 GGGGGGGAGGGTGG 742
 |||||

RESULT 11
 AJ437840
 LOCUS
 DEFINITION
 AJ437840 HapDlR2 Helianthus annuus cDNA clone HapDlR205H12, mRNA
 sequence.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Helianthus annuus (common sunflower)
 Helianthus annuus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; campanulids; Asterales; Asteraceae; Asteroideae;
 Heliantheae; Helianthus.
 1 (bases 1 to 511)
 Tamborindeguy,C., Ben,C., Liboz,T. and Gentzbbittel,L.
 Sequence evaluation of four specific cDNA libraries for
 developmental genomics of sunflower
 Mol. Genet. Genomics 271 (3), 367-375 (2004)
 Contact: Gentzbbittel L
 Laboratoire de Biotechnologie et Amélioration des Plantes
 Institut National Polytechnique de Toulouse - Ecole National
 Supérieure Agronomique de Toulouse
 IFR40, Pole de Biotechnologie Vegetale, 18 chemin de Borde Rouge,
 Auzeville, CASTANET TOLOSAN 31326, France.

FEATURES
 source
 1..511
 Location/Qualifiers
 /organism="Helianthus annuus"
 /mol_type="mRNA"
 /cultiivar="Emil"
 /db_xref="taxon:4232"
 /clone="HapDlR205H12"
 /tissue_type="hypocotyls"
 /cell_type="protoplasts"
 /dev_stage="1-5 days"
 /clone_lib="HapDlR2"

ORIGIN
 Query Match 12.0%; Score 39.2; DB 1; Length 511;
 Matches 75; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

Syngenta Arabidopsis Insertion Library (SAIL) lines are available through the Arabidopsis Biological Resource Center (ABRC). Sequences represent a pool of amplified genomic regions and not single contiguous sequences.
 Class: T-DNA tagged

FEATURES
 source
 1..936
 Location/Qualifiers
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /ecotype="Columbia"
 /db_xref="taxon:3702"
 /clone="SAIL_877_H06.v1"
 /clone_lib="SAIL Collection"
 /note="T-DNA left border sequences were isolated using a
 modified TAIL-PCR strategy"

ORIGIN
 Query Match 12.1%; Score 39.6; DB 9; Length 936;
 Best Local Similarity 56.0%; Pred. No. 4;
 Matches 75; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 188 AGTAGGCTGAGCACCGTGGCGGGCGGCGAGCGGTGGCGGTGGTGTCTTCTGGCGGA 247
 |||||
 DB 875 AGATGGGGGAGCCCGAGGGGGGGGGGAGCGGGGGGGGAGTGGGGGGGGGGGG 816
 |||||
 QY 248 GGTGCTGCTGATGATTAATTAAGTAGGCGGTCTTGAGACGCGGATGTCGAGGTGAG 307
 |||||
 DB 815 GGGGAGATGCTGTAGGCAAGAGATGGGACACAGGGGAGGTGGGTAGGGGGGGGG 756
 |||||
 QY 308 GTGTGGCAGGCTTG 321
 |||||
 DB 755 GGGGGGGAGGGTGG 742
 |||||

RESULT 11
 AJ437840
 LOCUS
 DEFINITION
 AJ437840 HapDlR2 Helianthus annuus cDNA clone HapDlR205H12, mRNA
 sequence.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Helianthus annuus (common sunflower)
 Helianthus annuus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; campanulids; Asterales; Asteraceae; Asteroideae;
 Heliantheae; Helianthus.
 1 (bases 1 to 511)
 Tamborindeguy,C., Ben,C., Liboz,T. and Gentzbbittel,L.
 Sequence evaluation of four specific cDNA libraries for
 developmental genomics of sunflower
 Mol. Genet. Genomics 271 (3), 367-375 (2004)
 Contact: Gentzbbittel L
 Laboratoire de Biotechnologie et Amélioration des Plantes
 Institut National Polytechnique de Toulouse - Ecole National
 Supérieure Agronomique de Toulouse
 IFR40, Pole de Biotechnologie Vegetale, 18 chemin de Borde Rouge,
 Auzeville, CASTANET TOLOSAN 31326, France.

FEATURES
 source
 1..511
 Location/Qualifiers
 /organism="Helianthus annuus"
 /mol_type="mRNA"
 /cultiivar="Emil"
 /db_xref="taxon:4232"
 /clone="HapDlR205H12"
 /tissue_type="hypocotyls"
 /cell_type="protoplasts"
 /dev_stage="1-5 days"
 /clone_lib="HapDlR2"

ORIGIN
 Query Match 12.0%; Score 39.2; DB 1; Length 511;
 Matches 75; Conservative 0; Mismatches 59; Indels 0; Gaps 0;


```
JOURNAL Unpublished (2003)
COMMENT Contact: Han Bin
National Center for Gene Research
Chinese Academy of Sciences
500# Cao Bao Road, Shanghai 200233, China
Email: bhan@ncgr.ac.cn
Clone requests: bhan@ncgr.ac.cn
This is rice cDNA est clone
Web site: http://www.ncgr.ac.cn.
Location/Qualifiers
1..847
/organism="Oryza sativa"
/mol_type="mRNA"
/db_xref="taxon:4530"
/clone="y658g03p5"
/clone_lib="Oryza sativa library (Han B)"

FEATURES
source
1..847
Query Match 12.0%; Score 39.2; DB 5; Length 847;
Best Local Similarity 53.7%; Pred. No. 5.1;
Matches 80; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 175 ACTCAGTCGCAAGGTAGGTCGACACCGTCGGCGGCGGACGGGTGGCGGTCTCGGGT 234
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
631 ACTCACACTCTCAAGTCCTCCAAACCAACAAGGGGGAGGGGGGGGTGTGGGGAGGGGG 690
QY 235 TGTTCCTGCGGAGGTGCTGCTGATCATGTATTAAGTAGCGGTCTTTGAGACCGCGGA 294
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 295 TGGTCGAGGTAGGTGTGGCGGCTTGAG 323
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
751 GGGTCGAGGAGGTGGAGACCGGAGAG 779

RESULT 15
CB685324/c
LOCUS 378 bp mRNA linear EST 09-APR-2003
DEFINITION OSJNEf15K06.f OSJNEf Oryza sativa (japonica cultivar-group) cDNA
clone OSJNEf15K06 5', mRNA sequence.
ACCESSION CB685324
VERSION CB685324.1 GI:29689049
SOURCE EST.
ORGANISM Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 378)
Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
Unpublished (2003)
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aaa cga cgg cca gtcg
BACKWARD: gga aac agc tat gac cat g
Plate: 15 row: K column: 06
Seq primer: gta aaa cga cgg cca gtcg.
Location/Qualifiers
1..378
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"

FEATURES
source
```

```
/clone="OSJNEf15K06"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSJNEf"
/note="Vector: pbluescript II KS +; Site_1: EcoRI; Site_2:
XhoI; Uninfected Control"

ORIGIN
Query Match 11.9%; Score 39; DB 6; Length 378;
Best Local Similarity 55.6%; Pred. No. 5.4;
Matches 75; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 111 CCGAGGGACCTGAGCGAGTCCGCATCGACCGGATCGGAAAACCTTCGAGAAAAGCGTCT 170
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
359 CCGAGGGCCCTTGGGGGGTCTCGTCCGCCACCTCCGGCGACGGCGGAGGAACCGGAAG 300
QY 171 AACCACTCACAGTCGCAAGGTAGGCTGAGCACCGTCGGCGGCGGCGGCTCG 230
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
299 GAGAAAGTCCGCGGCGCGCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCG 240
QY 231 GGGTTGTTTCTGGCG 245
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
239 CCGGGGTGGAAGCG 225

Search completed: July 14, 2005, 23:22:42
Job time : 3123.52 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:39:07 ; Search time 9289.17 Seconds
(without alignments)
6468.225 Million cell updates/sec

Title: US-09-482-682-32
Perfect score: 1240
Sequence: 1 ggaaccactctcttcgcacgttcacagtcgcaagatctc 1240

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_phi.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_scs.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1240	100.0	1240	6	BD268221 Adenoviru
2	1239	99.9	7231	6	BD268252 Adenoviru
3	1239	99.9	7960	6	BD268233 Adenoviru
4	1239	99.9	7989	6	BD268236 Adenoviru
5	1239	99.9	8383	6	BD268232 Adenoviru
6	1239	99.9	8484	6	BD268253 Adenoviru
7	1135	91.5	10332	6	A83180 Sequence 13
8	1135	91.5	10332	6	BD082846 Method an
9	1135	91.5	11570	14	AD5001
10	1135	91.5	31976	6	CQ854904 Sequence
11	1135	91.5	31976	6	CQ854905 Sequence
12	1135	91.5	32480	6	BD268216 Adenoviru
13	1135	91.5	32798	6	AR343138 Sequence
14	1135	91.5	32798	6	AX382187 Sequence
15	1135	91.5	32802	6	CQ854906 Sequence
16	1135	91.5	33007	12	AF323988
17	1135	91.5	33476	12	AY370909 Expressio
18	1135	91.5	33592	6	AX084504 Sequence
19	1135	91.5	33699	6	AX084506 Sequence

20	1135	91.5	33988	6	AX084517 Sequence
21	1135	91.5	34303	6	AR091536 Sequence
22	1135	91.5	34303	6	AR102229 Sequence
23	1135	91.5	34303	6	AR230727 Sequence
24	1135	91.5	34341	6	AX084505 Sequence
25	1135	91.5	34448	6	AX084507 Sequence
26	1135	91.5	34737	6	AX084518 Sequence
27	1135	91.5	35408	6	AR163568 Sequence
28	1135	91.5	35408	6	AR166442 Sequence
29	1135	91.5	35724	6	AX084516 Sequence
30	1135	91.5	35764	12	AY046510 Adenovira
31	1135	91.5	35871	6	AR403724 Sequence
32	1135	91.5	35934	14	AY339865 Human ade
33	1135	91.5	35935	6	AR091533 Sequence
34	1135	91.5	35935	6	AR102226 Sequence
35	1135	91.5	35935	6	AR116313 Sequence
36	1135	91.5	35935	6	CQ854907 Sequence
37	1135	91.5	35935	6	AR230724 Sequence
38	1135	91.5	35935	6	AX451988 Sequence
39	1135	91.5	35935	6	AX683770 Sequence
40	1135	91.5	35935	14	ADRCOMPGEN M73260 Mastadenovi
41	1135	91.5	35978	6	AR403723 Sequence
42	1135	91.5	36114	6	AX084519 Sequence
43	1135	91.5	36154	6	AX468857 Sequence
44	1135	91.5	36154	6	AX468865 Sequence
45	1135	91.5	36620	6	AR534337 Sequence

ALIGNMENTS

RESULT 1	BD268221	Adenovirus vector, packaging cell line, composition and method for production and use.	1240 bp	DNA	linear	PAT 17-JUL-2003
LOCUS	BD268221					
DEFINITION	BD268221.1	GI:33077989				
ACCESSION	BD268221					
VERSION	JP 2002534130-A/25.					
KEYWORDS	unidentified adenovirus					
SOURCE	unidentified adenovirus					
ORGANISM	Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.					
REFERENCE	1 (bases 1 to 1240)					
AUTHORS	Nemerow,G.R., Seggern,D.J.V., Hallenbeck,P.L., Stevenson,S.C. and Skripchenko,Y.					
TITLE	Adenovirus vector, packaging cell line, composition and method for production and use					
JOURNAL	Patent: JP 2002534130-A 25 15-OCT-2002;					
COMMENT	NOVARTIS AG, THE SCRIPPS RESEARCH INSTITUTE					
	OS Adenovirus					
	PN JP 2002534130-A/25					
	PD 15-OCT-2002					
	PF 14-JAN-2000 JP 2000593765					
	PR 14-JAN-1999 US 60/115920					
	PI GLEN ROBERT NEMEROW,DANIEL J VON SEGGERN,PAUL L HALLENBECK, PI SUSAN C STEVENSON,YELENA SKRIPCENKO					
	PC C12N15/09,A61K35/76,A61K48/00,A61P35/00,A61P43/00,A61P43/00,					
	PC C12N5/10,					
	PC C12N7/00,C12Q1/68,G01N33/53,G01N33/566,C12N15/00,C12N5/00 CC					
	Adenovirus vector, packaging cell line, composition and method					
	CC for					
	CC production and use					
	FH Key					
	FT source					
	FT Location/Qualifiers					
	1. .1240					
	/organism="Adenovirus"					
	Location/Qualifiers					
	1. .1240					
	/organism="unidentified adenovirus"					
	/mol_type="genomic DNA"					
	/db_xref="taxon:10535"					
FEATURES						
source						
ORIGIN						

```
Query Match      100.0%; Score 1240; DB 6; Length 1240;
Best Local Similarity 100.0%; Pred. No. 1e-237;
Matches 1240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATCCACTCTCTCCCGCATCGCTGTCTGCGAGGCGCAGCTGTGGGGTGAGTACTCCCT 60
DB 1 GGATCCACTCTCTCCCGCATCGCTGTCTGCGAGGCGCAGCTGTGGGGTGAGTACTCCCT 60

QY 61 CTGAAAAGCGGCATGACTTTCGCGCTAAGATTGTCAAGTTTCGAGGAGGATTT 120
DB 61 CTGAAAAGCGGCATGACTTTCGCGCTAAGATTGTCAAGTTTCGAGGAGGATTT 120

QY 121 GATATTACCTGGCCCGCGGTGATCCCTTTGAGGGTGCGCATCCATCTGGTCAGAAA 180
DB 121 GATATTACCTGGCCCGCGGTGATCCCTTTGAGGGTGCGCATCCATCTGGTCAGAAA 180

QY 181 GACAACTCTTTTGTGTCAAGCTTGGTGCGCAAAACGACCCGTAGAGGGCGTTGGACAGAA 240
DB 181 GACAACTCTTTTGTGTCAAGCTTGGTGCGCAAAACGACCCGTAGAGGGCGTTGGACAGAA 240

QY 241 CTGGCGATGAGACGCGAGGTTGGTTTGTTCGCGATCGCGCGCTCTCTGGCGCGCAT 300
DB 241 CTGGCGATGAGACGCGAGGTTGGTTTGTTCGCGATCGCGCGCTCTCTGGCGCGCAT 300

QY 301 GTTTAGCTGCAGTATTGCGCGCAACGACCGCCATTTCGGGAAAGAGCGGTGCGCTC 360
DB 301 GTTTAGCTGCAGTATTGCGCGCAACGACCGCCATTTCGGGAAAGAGCGGTGCGCTC 360

QY 361 GTGCGGCACAGGTGACGCGCAACCGCGTTGTGCGAGGGTGACAAGGTCAACGCTCGT 420
DB 361 GTGCGGCACAGGTGACGCGCAACCGCGTTGTGCGAGGGTGACAAGGTCAACGCTCGT 420

QY 421 GGTCTACTCTCCGCTAGCGCTGTTGGTTCAGCAGAGGCGCGCCCTTTCGCGCAGCA 480
DB 421 GGTCTACTCTCCGCTAGCGCTGTTGGTTCAGCAGAGGCGCGCCCTTTCGCGCAGCA 480

QY 481 GAATCGCGGTAGGGGTCTAGCTGCGTCTCGTCCGGGGGTCTGCGTCCACGGTAAAGAC 540
DB 481 GAATCGCGGTAGGGGTCTAGCTGCGTCTCGTCCGGGGGTCTGCGTCCACGGTAAAGAC 540

QY 541 CCGCGGCACAGCGCGCTGCAAGTAGTCTATCTTGCATCTCTTCAAGTCTTAGCGCTG 600
DB 541 CCGCGGCACAGCGCGCTGCAAGTAGTCTATCTTGCATCTCTTCAAGTCTTAGCGCTG 600

QY 601 CTGCCATCGCGCGCGCAAGCGCGCTGATCGGTGAGTGGGGAACCCCATGGCAT 660
DB 601 CTGCCATCGCGCGCGCAAGCGCGCTGATCGGTGAGTGGGGAACCCCATGGCAT 660

QY 661 GGGGTGGGTGAGCGGGGCGTACATGCCCAATGCTGTAACGTAAGGGGCTCTCT 720
DB 661 GGGGTGGGTGAGCGGGGCGTACATGCCCAATGCTGTAACGTAAGGGGCTCTCT 720

QY 721 GAGTATTCGAAGATGTAGGGTAGCATCTTCCACCGCGGATGCTGGCGCGCACGTAATC 780
DB 721 GAGTATTCGAAGATGTAGGGTAGCATCTTCCACCGCGGATGCTGGCGCGCACGTAATC 780

QY 781 GATATGTTCTGCGAGGAGCGAGAGGTGCGGAACGAGTTGCTACGGGCGGGTCTCT 840
DB 781 GATATGTTCTGCGAGGAGCGAGAGGTGCGGAACGAGTTGCTACGGGCGGGTCTCT 840

QY 841 TGCTCGGAAGACTATCTGCTGAAAGATGGCATGTGATGATGATGATGATGATGATGATG 900
DB 841 TGCTCGGAAGACTATCTGCTGAAAGATGGCATGTGATGATGATGATGATGATGATGATG 900

QY 901 GAAGACGTTGAAGCTGGCGTCTGTGAGACCTTACCGCGTCAACGACGAGGAGCGGTAGGA 960
DB 901 GAAGACGTTGAAGCTGGCGTCTGTGAGACCTTACCGCGTCAACGACGAGGAGCGGTAGGA 960

QY 961 GTGCGCAGCTTTGTGAACAGCTCGCGGTGACCTGACAGTCTAGGGCGCAGTACTCCAG 1020
DB 961 GTGCGCAGCTTTGTGAACAGCTCGCGGTGACCTGACAGTCTAGGGCGCAGTACTCCAG 1020

QY 1021 GGTTCCTTGATGATGCATCTTATCTCTGCTCCCTTTTTTTTCCACAGCTCGCGGTGAG 1080
```

```
DB 1021 GGTTCCTTGATGATGCATCTTATCTGCTCCCTTTTTTTTCCACAGCTCGCGGTGAG 1080

QY 1081 GACAAACTCTTCGCGGTCTTTCCAGTACTCTTGATCGGAAACCGCTCGGCTCCGAACG 1140
DB 1081 GACAAACTCTTCGCGGTCTTTCCAGTACTCTTGATCGGAAACCGCTCGGCTCCGAACG 1140

QY 1141 AGATCCGTAATCCGCGCGGAGGACCTGAGCGAGTCCGATCGACCGATCGGAAAC 1200
DB 1141 AGATCCGTAATCCGCGCGGAGGACCTGAGCGAGTCCGATCGACCGATCGGAAAC 1200

QY 1201 TCTCGAGAAAGCGCTTAACCAAGTCACAGTCGCAAGATCT 1240
DB 1201 TCTCGAGAAAGCGCTTAACCAAGTCACAGTCGCAAGATCT 1240

RESULT 2
BD268252 7231 bp DNA linear PAT 17-JUL-2003
LOCUS Adenovirus vector, packaging cell line, composition and method for
DEFINITION production and use.
ACCESSION BD268252
VERSION BD268252.1 GI:33078020
KEYWORDS JP 2002534130-A/56.
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 7231)
AUTHORS Nemerow,G.R., Seggern,D.J.V., Hallenbeck,P.L., Stevenson,S.C. and
Skripchenko,Y.
TITLE Adenovirus vector, packaging cell line, composition and method for
production and use
JOURNAL Patent: JP 2002534130-A 56 15-OCT-2002;
NOVARTIS AG,THE SCRIPPS RESEARCH INSTITUTE
COMMENT OS Artificial Sequence
PN JP 2002534130-A/56
PD 15-OCT-2002
PF 14-JAN-2000 JP 2000593765
PR 14-JAN-1999 US 60/115920
PI GLEN ROBERT NEMEROW,DANIEL J VON SEGGERN,PAUL L HALLENBECK, PI
SUSAN C STEVENSON,YELENA SKRIPCHENKO
PC C12N15/09,A61K35/76,A61K48/00,A61P35/00,A61P43/00,A61P43/00,
PC C12N5/10, C12Q1/68,G01N33/53,G01N33/566, C12N15/00, C12N5/00 CC
PC C12N7/00, C12Q1/68,G01N33/53,G01N33/566, C12N15/00, C12N5/00 CC
Description of Artificial Sequence: plasmid
FH Key Location/Qualifiers
FT source 1..7231
FT /organism='Artificial Sequence'.

FEATURES
source
1..7231
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

ORIGIN

Query Match 99.9%; Score 1239; DB 6; Length 7231;
Best Local Similarity 100.0%; Pred. No. 1.4e-237;
Matches 1239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATCCACTCTCTTCCCGCATCGCTGTCTGCGAGGCGCAGCTGTGGGGTGAGTACTCCCT 60
DB 849 GGATCCACTCTTCCCGCATCGCTGTCTGCGAGGCGCAGCTGTGGGGTGAGTACTCCCT 908

QY 61 CTGAAAAGCGGCATGACTTCTGCGCTAAGATTGTCAAGTTTCGAGGAGGATTT 120
DB 909 CTGAAAAGCGGCATGACTTCTGCGCTAAGATTGTCAAGTTTCGAGGAGGATTT 968

QY 121 GATATTACCTGGCCCGCGGTGATCCCTTTGAGGGTGCGCATCCATCTGGTCAGAAA 180
DB 969 GATATTACCTGGCCCGCGGTGATCCCTTTGAGGGTGCGCATCCATCTGGTCAGAAA 1028

QY 181 GACAACTCTTTTGTGTCAAGCTTGGTGCGCAAAACGACCCGTAGAGGGCGTTGGACAGCA 240
```

Db 1029 GACAATCTTTTGTGTTCAAGCTTGGTGGCAAAACGACCGTAGAGGGCGTTGGACAGCA 1088
Qy 241 CTTGGCGATGAGCGCAGGGTTTGGTTTTTGTGCGGATCGCGCGCTCTCTTGGCGCGCAT 300
Db 1089 CTTGGCGATGAGCGCAGGGTTTGGTTTTTGTGCGGATCGCGCGCTCTCTTGGCGCGCAT 1148
Qy 301 GTTTAGCTGACGTTATTCGCGCGCAACGACCGCCATTTCGGGAAAGACGGTGGTGGCTC 360
Db 1149 GTTTAGCTGACGTTATTCGCGCGCAACGACCGCCATTTCGGGAAAGACGGTGGTGGCTC 1208
Qy 361 GTGCGGACACGCTGACGCGCCAAACCGGGTTGTGCGGATGACAAAGTCAACGCTGT 420
Db 1209 GTGCGGACACGCTGACGCGCCAAACCGGGTTGTGCGGATGACAAAGTCAACGCTGT 1268
Qy 421 GGCTACCTCTCCGCTAGGCGCTGTTGTGTCAGCAGAGCGCGCCCTTTCGCGGAGCA 480
Db 1269 GGCTACCTCTCCGCTAGGCGCTGTTGTGTCAGCAGAGCGCGCCCTTTCGCGGAGCA 1328
Qy 481 GAATGGCGTAGGGGTCTAGCTGCGTCTGTCGCGGGGGTCTGCTCACGGTAAAGAC 540
Db 1329 GAATGGCGTAGGGGTCTAGCTGCGTCTGTCGCGGGGGTCTGCTCACGGTAAAGAC 1388
Qy 541 CCGGGGACGAGCGCGCTGCAAGTGTCTATCTTGCATCTTGCAGTCTAGCGCTG 600
Db 1389 CCGGGGACGAGCGCGCTGCAAGTGTCTATCTTGCATCTTGCAGTCTAGCGCTG 1448
Qy 601 CTGCCATGCGCGGGCGGCAAGCGCGCTGATGGTTGAGTGGGGGACCCCATGGCAT 660
Db 1449 CTGCCATGCGCGGGCGGCAAGCGCGCTGATGGTTGAGTGGGGGACCCCATGGCAT 1508
Qy 661 GGGTGGGTGAGCGGGAGCGTATGCGCCAAATGTCTGTAACGTAAGGGGTCTCT 720
Db 1509 GGGTGGGTGAGCGGGAGCGTATGCGCCAAATGTCTGTAACGTAAGGGGTCTCT 1568
Qy 721 GAGTATCCAAATATGATAGGTAGCTATCTTCCACCGCGGATGCTGGCGCGCACGTAATC 780
Db 1569 GAGTATCCAAATATGATAGGTAGCTATCTTCCACCGCGGATGCTGGCGCGCACGTAATC 1628
Qy 781 GTATAGTTCTGCGGAGGAGCGAGAGGTGCGGACCGAGTTGCTACGGCGGGCTGCTC 840
Db 1629 GTATAGTTCTGCGGAGGAGCGAGAGGTGCGGACCGAGTTGCTACGGCGGGCTGCTC 1688
Qy 841 TGCTCGGAAGACTATCTGCTGAAAGTGGCATGTGAGTTGGATGATATGTTTGGACGCTG 900
Db 1689 TGCTCGGAAGACTATCTGCTGAAAGTGGCATGTGAGTTGGATGATATGTTTGGACGCTG 1748
Qy 901 GAAGAGCTTGAAGCTGGGCTGTGAGACCTACCGGTCACGACGAGGAGCGTAGGA 960
Db 1749 GAAGAGCTTGAAGCTGGGCTGTGAGACCTACCGGTCACGACGAGGAGCGTAGGA 1808
Qy 961 GTGCGGACGTTGTTGACGAGCTCGGCGGTGACCTGACGCTTAGGGCGCAGTAGTCCAG 1020
Db 1809 GTGCGGACGTTGTTGACGAGCTCGGCGGTGACCTGACGCTTAGGGCGCAGTAGTCCAG 1868
Qy 1021 GGTTCCTTTGATGATGATCTATCTGTCCTCTTTTTCACAGCTCGCGGTGAG 1080
Db 1869 GGTTCCTTTGATGATGATCTATCTGTCCTCTTTTTCACAGCTCGCGGTGAG 1928
Qy 1081 GACAACTCTTCGCGGCTTTTCCAGTACTCTTTGATGCGGAAACCGCTCGGCTCCGACG 1140
Db 1929 GACAACTCTTCGCGGCTTTTCCAGTACTCTTTGATGCGGAAACCGCTCGGCTCCGACG 1988
Qy 1141 AGATCGTACTCGCGCGGAGGACCTGAGGAGTCCGCATCGAGCGGATCGGAAACC 1200
Db 1989 AGATCGTACTCGCGCGGAGGACCTGAGGAGTCCGCATCGAGCGGATCGGAAACC 2048
Qy 1201 TCTCGAAGAGGGCTCTAACCAAGTACAGTTCGCAAGATC 1239
Db 2049 TCTCGAAGAGGGCTCTAACCAAGTACAGTTCGCAAGATC 2087

LOCUS BD268233 7960 bp DNA linear PAT 17-JUL-2003
DEFINITION Adenovirus vector, packaging cell line, composition and method for
production and use.
ACCESSION BD268233
VERSION BD268233.1 GI:33078001
KEYWORDS JP 2002534130-A/37
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 7960)
AUTHORS Nemerow,G.R., Seggern,D.J.V., Hallenbeck,P.L., Stevenson,S.C. and
Skripchenko,Y.
TITLE Adenovirus vector, packaging cell line, composition and method for
production and use
JOURNAL Patent: JP 2002534130-A 37 15-OCT-2002;
NOVARTIS AG,THE SCRIPPS RESEARCH INSTITUTE
COMMENT OS Artificial Sequence
PN JP 2002534130-A/37
PD 15-OCT-2002 JP 2000593765
PF 14-JAN-2000 JP 2000593765
PR 14-JAN-1999 US 60/115920
PI GLEN ROBERT NEMEROW,DANIEL J VON SEGGERN,PAUL L HALLENBECK, PI
SUSAN C STEVENSON,YELENA SKRIPCENKO
PC C12N15/09,A61K35/76,A61K48/00,A61P35/00,A61P43/00,A61P43/00,
PC C12N5/10,
PC C12N7/00,C12Q1/68,G01N33/53,G01N33/566,C12N15/00,C12N5/00 CC
Description of Artificial Sequence: plasmid
FH Key Location/Qualifiers
FT source 1..7960
/organism='Artificial Sequence'.
FEATURES
source
1..7960
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
ORIGIN
Query Match 99.9%; Score 1239; DB 6; Length 7960;
Best Local Similarity 100.0%; Pred. No. 1.4e-237;
Matches 1239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGATCCACTCTCTCCGCATCGTCTGCGAGGGCCAGCTGTGGGGTGAAGTACTCCCT 60
Db 929 GGATCCACTCTCTCCGCATCGTCTGCGAGGGCCAGCTGTGGGGTGAAGTACTCCCT 988
Qy 61 CTGAAAACGGCGATGACTTCTGCGCTAAGATTGTTCAGTTTCCAAAACGAGAGGATTT 120
Db 989 CTGAAAACGGCGATGACTTCTGCGCTAAGATTGTTCAGTTTCCAAAACGAGAGGATTT 1048
Qy 121 GATATTACCTTGGCCCGGCGTGATGCTTTGAGGGTGGCCGCATCCATCTGGTCAGAAAA 180
Db 1049 GATATTACCTTGGCCCGGCGTGATGCTTTGAGGGTGGCCGCATCCATCTGGTCAGAAAA 1108
Qy 181 GACAACTCTTTTGTGTCAGCTTGGTGGCAAAACGACCGGTAGAGGGGTGGACAGCA 240
Db 1109 GACAACTCTTTTGTGTCAGCTTGGTGGCAAAACGACCGGTAGAGGGGTGGACAGCA 1168
Qy 241 CTTGCGATGAGCGCAGGGTTTGGTTTTTGTGCGGATCGGCGCTCTCTTGGCGCGCAT 300
Db 1169 CTTGCGATGAGCGCAGGGTTTGGTTTTTGTGCGGATCGGCGCTCTCTTGGCGCGCAT 1228
Qy 301 GTTTAGCTGACGTTATTCGCGCAACGACCGCATTCGGGAAAGACGGTGGTGGCTC 360
Db 1229 GTTTAGCTGACGTTATTCGCGCAACGACCGCATTCGGGAAAGACGGTGGTGGCTC 1288
Qy 361 GTCGGGACCAAGGTGACGCGCAACCGCGGTGTGCAAGGTGCAAGCTCAACGCTGT 420
Db 1289 GTCGGGACCAAGGTGACGCGCAACCGCGGTGTGCAAGGTGCAAGCTCAACGCTGT 1348
Qy 421 GGCTACCTCTCGCGTAGCGCTCGTGGTCCAGCAGAGCGCGGCTTGGCGGAGCA 480
Db 1349 GGCTACCTCTCGCGTAGCGCTCGTGGTCCAGCAGAGCGCGGCTTGGCGGAGCA 1408

```
QY 481 GAATGCGGTAGGGGCTCTAGCTGCTCTCGTCCGGGGGCTCTGCTCCACGGTAAAGAC 540
Db 1409 GAATGCGGTAGGGGCTCTAGCTGCTCTCGTCCGGGGGCTCTGCTCCACGGTAAAGAC 1468
QY 541 CCCGGGCGACGAGCGCGCTCGAAGTAGTCTATCTTTGTCATCTTTCGAAGTCTTAGCGCCTG 600
Db 1469 CCCGGGCGACGAGCGCGCTCGAAGTAGTCTATCTTTGTCATCTTTCGAAGTCTTAGCGCCTG 1528
QY 601 CTGCCATGCGGGGCGGCAAGCGCGCTCGTATGCTGCTGAGTGGGGGACCCCATGGCAT 660
Db 1529 CTGCCATGCGGGGCGGCAAGCGCGCTCGTATGCTGCTGAGTGGGGGACCCCATGGCAT 1588
QY 661 GGGGTGGGTGACGCGGGGCTTACATCGCCAAATGTCGTAACGTAAGGGGCTCTCT 720
Db 1589 GGGGTGGGTGACGCGGGGCTTACATCGCCAAATGTCGTAACGTAAGGGGCTCTCT 1648
QY 721 GAGTATTCGAAGATATGATAGGTAGCATCTTCCACCGCGGATGCTGGCGCGCACGTAATC 780
Db 1649 GAGTATTCGAAGATATGATAGGTAGCATCTTCCACCGCGGATGCTGGCGCGCACGTAATC 1708
QY 781 GTATAGTTCTGCGAGGAGCGAGAGGTGCGGACCGAGTTGCTTACGGGCGGGTCTCTC 840
Db 1709 GTATAGTTCTGCGAGGAGCGAGAGGTGCGGACCGAGTTGCTTACGGGCGGGTCTCTC 1768
QY 841 TGCTCGGAAGACTATCTGCTCAAGATGCGATGCGATGCTGATGCTGATGATATGTTGGACGCTG 900
Db 1769 TGCTCGGAAGACTATCTGCTCAAGATGCGATGCGATGCTGATGCTGATGATATGTTGGACGCTG 1828
QY 901 GAAGACGTTGAAGCTGGGCTCTGTGAGACCTTACCGCGTCACGACGAAGGAGCGGTAGGA 960
Db 1829 GAAGACGTTGAAGCTGGGCTCTGTGAGACCTTACCGCGTCACGACGAAGGAGCGGTAGGA 1888
QY 961 GTGCGGAGCTTTGTGACAGCTCGGCGGTGACCTGACAGTCTAGGGCGCGAGTAGTCAG 1020
Db 1889 GTGCGGAGCTTTGTGACAGCTCGGCGGTGACCTGACAGTCTAGGGCGCGAGTAGTCAG 1948
QY 1021 GGTTCCTTGATGATCTATCTATCTTCCAGTCTTTCGATCGGAAACCGCTCCGGAACG 1140
Db 1949 GGTTCCTTGATGATCTATCTTTCGATCTTTCGATCGGAAACCGCTCCGGAACG 2008
QY 1081 GACAAACTCTTCGCGGTCTTTCCAGTACTTTCGATCGGAAACCGCTCCGGAACG 1140
Db 2009 GACAAACTCTTCGCGGTCTTTCCAGTACTTTCGATCGGAAACCGCTCCGGAACG 2068
QY 1141 AGATCCGTACTCCGCGCGGAGGACCTGAGCGAGTCCGCGATCGACCGGATCGGAAACG 1200
Db 2069 AGATCCGTACTCCGCGCGGAGGACCTGAGCGAGTCCGCGATCGACCGGATCGGAAACG 2128
QY 1201 TCTCGAGAAAGCGCTTACCAAGTCCAGTCCGGAAGTC 1239
Db 2129 TCTCGAGAAAGCGCTTACCAAGTCCAGTCCGGAAGTC 2167

RESULT 4
LOCUS BD268236
DEFINITION Adenovirus vector, packaging cell line, composition and method for production and use.
ACCESSION BD268236
VERSION BD268236.1 GI:33078004
KEYWORDS JP 2002534130-A/40
SOURCE synthetic construct
ORGANISM
REFERENCE
AUTHORS Nemerow,G.R., Seggern,D.J.V., Hallenbeck,P.L., Stevenson,S.C. and Skripchenko,Y.
TITLE Adenovirus vector, packaging cell line, composition and method for production and use
JOURNAL. Patent: JP 2002534130-A 40 15-OCT-2002; NOVARTIS AG,THE SCRIPPS RESEARCH INSTITUTE
COMMENT OS Artificial Sequence
PN JP 2002534130-A/40
```

```
PD 15-OCT-2002
PF 14-JAN-2000 JP 2000593765
PI 14-JAN-1999 US 60/115920
PR GLEN ROBERT NEMEROW,DANIEL J VON SEGGERN,PAUL L HALLENBECK, PI
PC SUSAN C STEVENSON,YELENA SKRIPCHENKO
PC C12N15/09,A61K35/76,A61K48/00,A61P35/00,A61P43/00,A61P43/00,
PC C12N5/10,
PC C12N7/00,C12Q1/68,G01N33/53,G01N33/566,C12N15/00,C12N5/00 CC
Description of Artificial Sequence: plasmid
FT Key Location/Qualifiers
FT source 1..7989
FT /organism='Artificial Sequence'.

FEATURES
source
1..7989
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

ORIGIN
Query Match 99.9%; Score 1239; DB 6; Length 7989;
Best Local Similarity 100.0%; Pred. No. 1.4e-237;
Matches 1239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATCCACTCTCTTCCGATCGCTGTCTGCGAGGGCCAGCTGTTGGGGTGAGTACTCCCT 60
Db 929 GGATCCACTCTCTTCCGATCGCTGTCTGCGAGGGCCAGCTGTTGGGGTGAGTACTCCCT 988
QY 61 CTGAAAACGCGGCATGACTTCTGCGCTAAGATTCTCAGTTTCCAAAAACGAGGAGGATTT 120
Db 989 CTGAAAACGCGGCATGACTTCTGCGCTAAGATTCTCAGTTTCCAAAAACGAGGAGGATTT 1048
QY 121 GATATTACCTGGCGCGCGGTGATGCTTTCAGGGTGGCGGCATCCATCTGTCAGAAAA 180
Db 1049 GATATTACCTGGCGCGCGGTGATGCTTTCAGGGTGGCGGCATCCATCTGTCAGAAAA 1108
QY 181 GACAACTCTTTTGTGTGCAAGCTTGTGGCAACGACCCGTAGAGGGGCTTGGACAGCAA 240
Db 1109 GACAACTCTTTTGTGTGCAAGCTTGTGGCAACGACCCGTAGAGGGGCTTGGACAGCAA 1168
QY 241 CTTGCGCATGAGAGCGAGGGTTTGGTTTGTTCGCGATCGCGCGCTCTTTGGCGCGCAT 300
Db 1169 CTTGCGCATGAGAGCGAGGGTTTGGTTTGTTCGCGATCGCGCGCTCTTTGGCGCGCAT 1228
QY 301 GTTTAGTGCACGTATTTCGCGCGCAACGACCGCCATTCGGAAGACGGTGGCGCTC 360
Db 1229 GTTTAGTGCACGTATTTCGCGCGCAACGACCGCCATTCGGAAGACGGTGGCGCTC 1288
QY 361 GTGCGGCAACGAGGTGCAACGCGCTTGTGAGGGTGACAAAGGTCAAAGCTCGGT 420
Db 1289 GTGCGGCAACGAGGTGCAACGCGCTTGTGAGGGTGACAAAGGTCAAAGCTCGGT 1348
QY 421 GGCTACTCTCTCCGCTAGGGCTCGTTGGTCCAGAGCGCGCGCTTTGGCGAGCA 480
Db 1349 GGCTACTCTCTCCGCTAGGGCTCGTTGGTCCAGAGCGCGCGCTTTGGCGAGCA 1408
QY 481 GAATGGCGGTAGGGGCTTAGCTCGCTCTGTCGGGGGGTCTGGGTCCAGGTAAAGAC 540
Db 1409 GAATGGCGGTAGGGGCTTAGCTCGCTCTGTCGGGGGGTCTGGGTCCAGGTAAAGAC 1468
QY 541 CCCGGGCGACGAGCGCGCTCGTCAAGTAGTCTATCTTTCATCTTTCGAAGTCTTAGCGCCTG 600
Db 1469 CCCGGGCGACGAGCGCGCTCGTCAAGTAGTCTATCTTTCATCTTTCGAAGTCTTAGCGCCTG 1528
QY 601 CTGCCATGCGGGGCGGCAAGCGCGCTCGTATGCTGCTGAGTGGGGGACCCCATGGCAT 660
Db 1529 CTGCCATGCGGGGCGGCAAGCGCGCTCGTATGCTGCTGAGTGGGGGACCCCATGGCAT 1588
QY 661 GGGGTGGGTGACGCGGGGCTTACATCGCCAAATGTCGTAACGTAAGGGGCTCTCT 720
Db 1589 GGGGTGGGTGACGCGGGGCTTACATCGCCAAATGTCGTAACGTAAGGGGCTCTCT 1648
QY 721 GAGTATTCGAAGATATGATAGGTAGCATCTTCCACCGCGGATGCTGGCGCGCACGTAATC 780
Db 1649 GAGTATTCGAAGATATGATAGGTAGCATCTTCCACCGCGGATGCTGGCGCGCACGTAATC 1708
```

Db 1649 GAGTATTCAGAGATATGTAGGTAGCATCTTCCACCGCGATGCTGCGCGCACCACTAATC 1708
Qy 781 GTATAGTTCTGTCGAGGAGAGAGAGGTGCGGACCGAGTTGCTACCGGCGGCTGCTC 840
Db 1709 GTATAGTTCTGTCGAGGAGAGAGAGGTGCGGACCGAGTTGCTACCGGCGGCTGCTC 1768
Qy 841 TCTCTGGAAGACTATCTGCTCTGAAAGATGCGATGTGAGTTGGATGATATGTTGGACGCTG 900
Db 1769 TCTCTGGAAGACTATCTGCTCTGAAAGATGCGATGTGAGTTGGATGATATGTTGGACGCTG 1828
Qy 901 GAAGACGTTGAAGCTGGCTCTGTGAGACCTACCGCTCACCGACGAGGAGGCTGAGGA 960
Db 1829 GAAGACGTTGAAGCTGGCTCTGTGAGACCTACCGCTCACCGACGAGGAGGCTGAGGA 1888
Qy 961 GTCTGGCGACGTTGTTGACAGCTCGCGGTGACCTGCACGCTAGGCGGCGAGTAGTCCAG 1020
Db 1889 GTCTGGCGACGTTGTTGACAGCTCGCGGTGACCTGCACGCTAGGCGGCGAGTAGTCCAG 1948
Qy 1021 GGTTCCTTGATGATGTCTATCTTCCAGTACTCTTGGATCGGAACCGCTCGGCTCCGAACG 1140
Db 1949 GGTTCCTTGATGATGTCTATCTTCCAGTACTCTTGGATCGGAACCGCTCGGCTCCGAACG 2068
Qy 1081 GACAACTCTTCGCGGTCTTCCAGTACTCTTGGATCGGAACCGCTCGGCTCCGAACG 1140
Db 2009 GACAACTCTTCGCGGTCTTCCAGTACTCTTGGATCGGAACCGCTCGGCTCCGAACG 2068
Qy 1141 AGATCCGCTACTCGCGCGGAGGACCTGAGCGAGTCCGCATCGACCGGATCGGAACCG 1200
Db 2069 AGATCCGCTACTCGCGCGGAGGACCTGAGCGAGTCCGCATCGACCGGATCGGAACCG 2128
Qy 1201 TCTCGAGAAAGCGCTTACCAAGTCCAGTCCCAAGATC 1239
Db 2129 TCTCGAGAAAGCGCTTACCAAGTCCAGTCCCAAGATC 2167

RESULT 5

BD268232
LOCUS Adenovirus vector, packaging cell line, composition and method for production and use. 8383 bp DNA linear PAT 17-JUL-2003
DEFINITION
ACCESSION BD268232
VERSION BD268232.1 GI:33078000
KEYWORDS JP 2002534130-A/36.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 8383)
AUTHORS Nemerow,G.R., Seggern,D.J.V., Hallenbeck,P.L., Stevenson,S.C. and Skripchenko,Y.
TITLE Adenovirus vector, packaging cell line, composition and method for production and use.
JOURNAL Patent: JP 2002534130-A 36 15-OCT-2002;
COMMENT NOVARTIS AG,THE SCRIPPS RESEARCH INSTITUTE
OS Artificial Sequence
PN JP 2002534130-A/36
PD 15-OCT-2002
PF 14-JAN-2000 JP 2000593765
PR 14-JAN-1999 US 60/115920
PI GLEN ROBERT NEMEROW,DANIEL J VON SEGGERN,PAUL L HALLENBECK, PI
PC SUSAN C STEVENSON,YELENA SKRIPCHENKO
PC C12N15/09,A61K35/76,A61K48/00,A61P35/00,A61P43/00,A61P43/00,
PC C12N5/10.
PC C12N7/00,C12Q1/68,C01N33/53,C01N33/566,C12N15/00,C12N5/00 CC
Description of Artificial Sequence: plasmid
FH Key Location/Qualifiers
FT source 1..8383
FT Location/Qualifiers
FEATURES
source 1..8383
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

ORIGIN

Query Match 99.9%; Score 1239; DB 6; Length 8383;
Best Local Similarity 100.0%; Pred. No. 1.4e-237;
Matches 1239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGATCCACTCTCTCCGATCGCTGTCTGCGAGGGCCAGCTGTTGGGGTGAGTATCCCT 60
Db 907 GGATCCACTCTCTCCGATCGCTGTCTGCGAGGGCCAGCTGTTGGGGTGAGTATCCCT 966
Qy 61 CTGAAAACGGGCGATGACTTCTGCGCTAAGATTGTCACTTTCCAAAAACGAGGAGGATTT 120
Db 967 CTGAAAACGGGCGATGACTTCTGCGCTAAGATTGTCACTTTCCAAAAACGAGGAGGATTT 1026
Qy 121 GATATTACCTTGGCCGCGGTATGCTTTGAGGGTGGCCGCATCCATCTGCTCAGAAAA 180
Db 1027 GATATTACCTTGGCCGCGGTATGCTTTGAGGGTGGCCGCATCCATCTGCTCAGAAAA 1086
Qy 181 GACATCTTTTGTGTCAAGCTTGGTGGCAACGACCCGTAGAGGGGCTTGGACAGCA 240
Db 1087 GACATCTTTTGTGTCAAGCTTGGTGGCAACGACCCGTAGAGGGGCTTGGACAGCA 1146
Qy 241 CTTCGCGATGAGCGCAGGGTTTGGTTTGTCTGCGATCGCGGCTCTCTTGGCCGCGAT 300
Db 1147 CTTCGCGATGAGCGCAGGGTTTGGTTTGTCTGCGATCGCGGCTCTCTTGGCCGCGAT 1206
Qy 301 GTTTAGCTGCACTATTTCGCGCAACGACCCGCAATTCGGGAAAGACGGTGGTGCCTC 360
Db 1207 GTTTAGCTGCACTATTTCGCGCAACGACCCGCAATTCGGGAAAGACGGTGGTGCCTC 1266
Qy 361 GTTCGGGACCAAGGTGACCGCGCAACCGCGTTGTGAGGGTGACAAGGTCAACGCTGCT 420
Db 1267 GTTCGGGACCAAGGTGACCGCGCAACCGCGTTGTGAGGGTGACAAGGTCAACGCTGCT 1326
Qy 421 GGTACCTCTCCGCGTAGCGCTGTTGCTCCAGAGAGGGCGCGCTCTGCGCGAGCA 480
Db 1327 GGTACCTCTCCGCGTAGCGCTGTTGCTCCAGAGAGGGCGCGCTCTGCGCGAGCA 1386
Qy 481 GAAATGGCGGTAGGGGCTTAGCTGCGTCTCGTCCGGGGGCTGCTGCTCCACGGTAAAGAC 540
Db 1387 GAAATGGCGGTAGGGGCTTAGCTGCGTCTCGTCCGGGGGCTGCTGCTCCACGGTAAAGAC 1446
Qy 541 CCCGGGACGAGCGCGCTGCAAGTAGCTATCTTTGCAATCTTCCAAAGTCTAGCGCTG 600
Db 1447 CCCGGGACGAGCGCGCTGCAAGTAGCTATCTTTGCAATCTTCCAAAGTCTAGCGCTG 1506
Qy 601 CTGCGCATCGCGCGGCGCAAGCGCGCTGCTGTTGAGTGGGGGACCCCATGGCAT 660
Db 1507 CTGCGCATCGCGCGGCGCAAGCGCGCTGCTGTTGAGTGGGGGACCCCATGGCAT 1566
Qy 661 GGGGTGGGTGAGCGCGAGGCGTACATCCCGCAATGTCTGTAACGTAGAGGGGCTCTCT 720
Db 1567 GGGGTGGGTGAGCGGAGGCGTACATCCCGCAATGTCTGTAACGTAGAGGGGCTCTCT 1626
Qy 721 GAGTATTCGAAGATATGATGAGGTAGCATCTTCCACCGCGGATGTGCGCGCACGTAATC 780
Db 1627 GAGTATTCGAAGATATGATGAGGTAGCATCTTCCACCGCGGATGTGCGCGCACGTAATC 1686
Qy 781 GTATAGTTCTGTCGAGGAGCGAGGAGTCCGACCGAGGTTGCTACCGGCGGGCTGCTC 840
Db 1687 GTATAGTTCTGTCGAGGAGCGAGGAGTCCGACCGAGGTTGCTACCGGCGGGCTGCTC 1746
Qy 841 TGCTCGGAAGACTATCTCCCTGAAGATGGCATGTGAGTTGGATGATATGTTTGAACGCTG 900
Db 1747 TGCTCGGAAGACTATCTCCCTGAAGATGGCATGTGAGTTGGATGATATGTTTGAACGCTG 1806
Qy 901 GAAGACGTTGAAGCTGGCTCTGTGAGACCTACCGCGTCCAGCAAGAGGAGGCTAGGA 960
Db 1807 GAAGACGTTGAAGCTGGCTCTGTGAGACCTACCGCGTCCAGCAAGAGGAGGCTAGGA 1866
Qy 961 GTTCGGGACGTTGTTGACCAAGCTCGGCGGTGACCTGACGCTAGGCGCGAGTAGTCCAG 1020
Db 1867 GTTCGGGACGTTGTTGACCAAGCTCGGCGGTGACCTGACGCTAGGCGCGAGTAGTCCAG 1926

QY 1021 GGTTCCTTGATGATCATCTTATCCTGTCCTCTTTTTCACACGCTCGCGTTGAG 1080
Db |||||
1927 GGTTCCTTGATGATCATCTTATCCTGTCCTCTTTTTCACACGCTCGCGTTGAG 1986
QY 1081 GACAACTCTTCGCGGTCTTTCCAGTACTCTTGAGTCGGAACCCGTCGCGCTCCGAACG 1140
Db |||||
1987 GACAACTCTTCGCGGTCTTTCCAGTACTCTTGAGTCGGAACCCGTCGCGCTCCGAACG 2046
QY 1141 AGATCCGTACTCCGCGCGGAGGACCTGAGGAGTCGCGATCGACCGGATCGGAACACC 1200
Db |||||
2047 AGATCCGTACTCCGCGCGGAGGACCTGAGGAGTCGCGATCGACCGGATCGGAACACC 2106
QY 1201 TCTCAGAAAGCGCTAACAGTCACAGTCGCAAGATC 1239
Db |||||
2107 TCTCAGAAAGCGCTAACAGTCACAGTCGCAAGATC 2145

RESULT 6

BD268253 8484 bp DNA linear PAT 17-JUL-2003
LOCUS Adenovirus vector, packaging cell line, composition and method for
DEFINITION production and use.
ACCESSION BD268253
VERSION BD268253.1 GI:33078021
KEYWORDS JP 2002534130-A/57.
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 8484)
AUTHORS Nemerow,G.R., Seggern,D.J.V., Hallenbeck,P.L., Stevenson,S.C. and
Sripichenko,Y.

TITLE Adenovirus vector, packaging cell line, composition and method for
production and use

JOURNAL Patent: JP 2002534130-A 57 15-OCT-2002;

NOVARTIS AG,THE SCRIPPS RESEARCH INSTITUTE

COMMENT OS Artificial Sequence

PN JP 2002534130-A/57

PD 15-OCT-2002

PF 14-JAN-2000 JP 2000593765

PI 14-JAN-1999 US 60/115920

PI GLEN ROBERT NEMEROW,DANIEL J VON SEGGERN,PAUL L HALLENBECK, PI

SUSAN C STEVENSON,YELINA SKRIPCHENKO

PC C12N15/09,A61K35/76,A61K48/00,A61P35/00,A61P43/00,A61P43/00,

PC C12N5/10,

PC C12N7/00,C12Q1/68,G01N33/53,G01N33/566,C12N15/00,C12N5/00 CC

Description of Artificial Sequence: plasmid

FH Key Location/Qualifiers

FT source 1..8484

FT /organism='Artificial Sequence'.

Location/Qualifiers

source 1..8484

/organism='synthetic construct'

/mol_type='genomic DNA'

/db_xref='taxon:32630'

ORIGIN

Query Match 99.9%; Score 1239; DB 6; Length 8484;

Best Local Similarity 100.0%; Pred. No. 1.4e-237;

Matches 1239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGATCCACTCTTCGCGATCGCTCTCTCGGAGGCCAGCTGTTGGGTGAGTACTCCCT 60
Db |||||
849 GGATCCACTCTTCGCGATCGCTCTCTCGGAGGCCAGCTGTTGGGTGAGTACTCCCT 908

QY 61 CTGAAAAGCGGCATGACTTCTCGCGTAAGATTGTTCAGTTTCCAAAAACGAGGAGATT 120
Db |||||

QY 909 CTGAAAAGCGGCATGACTTCTCGCGTAAGATTGTTCAGTTTCCAAAAACGAGGAGATT 968
Db |||||

QY 121 GATATTACCTGGCCGCGGTGATGCTTTGAGGGTGGCCGATCCATCTGGTTCAGAAA 180
Db |||||

QY 969 GATATTACCTGGCCGCGGTGATGCTTTGAGGGTGGCCGATCCATCTGGTTCAGAAA 1028
Db |||||

QY 181 GACAATCTTTTGTGTCAAGCTTGGTGGCAAAACGACCCGTAGAGGGCGTTGGACAGCAA 240
Db |||||

RESULT 7

Db |||||
1029 GACAATCTTTTGTGTCAAGCTTGGTGGCAAAACGACCCGTAGAGGGCGTTGGACAGCAA 1088
QY 241 CTTGGCGATGAGCGCAGGGGTTGGTTTTTGTTCGCGATCGCGCGCTCTTTGGCCGCGAT 300
Db |||||
1089 CTTGGCGATGAGCGCAGGGGTTGGTTTTTGTTCGCGATCGCGCGCTCTTTGGCCGCGAT 1148
QY 301 GTTTAGCTGACATTTTCGCGCGCAACGCCACCGCATTCGGGAAAGACGGTGGTCCGCTC 360
Db |||||
1149 GTTTAGCTGACATTTTCGCGCGCAACGCCACCGCATTCGGGAAAGACGGTGGTCCGCTC 1208
QY 361 GTCGGGCACCAAGGTGACCGCGCTTGTGAGGGTGCACAAAGGTCAACGCTGGT 420
Db |||||
1209 GTCGGGCACCAAGGTGACCGCGCTTGTGAGGGTGCACAAAGGTCAACGCTGGT 1268
QY 421 GGCTACCTCTCCGCTAGGCGCTCGTTGGTCCAGCAGAGCGCGCCCTTTGCCGAGACA 480
Db |||||
1269 GGCTACCTCTCCGCTAGGCGCTCGTTGGTCCAGCAGAGCGCGCCCTTTGCCGAGACA 1328
QY 481 GAATGGCGGTAGGGGTTAGCTGCGTCTCGTCCGGGGGTTCTGCTCCACGGTAAAGAC 540
Db |||||
1329 GAATGGCGGTAGGGGTTAGCTGCGTCTCGTCCGGGGGTTCTGCTCCACGGTAAAGAC 1388
QY 541 CCGCGCAGCAGGCGCGCTCGAAGTAGTCTATCTTGCATCTTTGCAAGTCTAGCGCTG 600
Db |||||
1389 CCGCGCAGCAGGCGCGCTCGAAGTAGTCTATCTTGCATCTTTGCAAGTCTAGCGCTG 1448
QY 601 CTGCCATCGCGCGCGCAAGCGCGCTCGTATGGTGTGAGTGGGGACCCCATGGCAT 660
Db |||||
1449 CTGCCATCGCGCGCGCAAGCGCGCTCGTATGGTGTGAGTGGGGACCCCATGGCAT 1508
QY 661 GGGTGGGTGAGCGCGAGGGGTACATCCGCAATGTCTGTAACGTAGAGGGGCTCTCT 720
Db |||||
1509 GGGTGGGTGAGCGCGAGGGGTACATCCGCAATGTCTGTAACGTAGAGGGGCTCTCT 1568
QY 721 GAGTATTTCCAAGATATGTAGGGTAGCATCTTCCACCGCGATGTGGCGCGACAGTAATC 780
Db |||||
1569 GAGTATTTCCAAGATATGTAGGGTAGCATCTTCCACCGCGATGTGGCGCGACAGTAATC 1628
QY 781 GTATAGTTCTGCGAGGGAGCGAGAGGTGGGACCGAGGTTGTACGGGCGGCTGCTC 840
Db |||||
1629 GTATAGTTCTGCGAGGGAGCGAGAGGTGGGACCGAGGTTGTACGGGCGGCTGCTC 1688
QY 841 TGCTCGAAGACTATCTGCTGAAGATGGCATGTGAGTGGATGATATGTTGGACCGCTG 900
Db |||||
1689 TGCTCGAAGACTATCTGCTGAAGATGGCATGTGAGTGGATGATATGTTGGACCGCTG 1748
QY 901 GAAGACGTTGAAGCTGGCGCTCTGTGAGACCTTACCGCGTCACGCACGAAGAGGGCTAGGA 960
Db |||||
1749 GAAGACGTTGAAGCTGGCGCTCTGTGAGACCTTACCGCGTCACGCACGAAGAGGGCTAGGA 1808
QY 961 GTCGCGAGCTGTTGTGACCAAGCTCGGCGTGACCTGCA CGTCTAGGGCGCAGTAGTCCAG 1020
Db |||||
1809 GTCGCGAGCTGTTGTGACCAAGCTCGGCGTGACCTGCA CGTCTAGGGCGCAGTAGTCCAG 1868
QY 1021 GGTTCCTTGATGATGTCATCTTATCCTGTCCTCTTTTTCACACGCTCGCGTTGAG 1080
Db |||||
1869 GGTTCCTTGATGATGTCATCTTATCCTGTCCTCTTTTTCACACGCTCGCGTTGAG 1928
QY 1081 GACAACTCTTCGCGGTCTTTTCAGTACTCTTTGATCGGAAACCCGTCGCGCTCCGAACG 1140
Db |||||
1929 GACAACTCTTCGCGGTCTTTTCAGTACTCTTTGATCGGAAACCCGTCGCGCTCCGAACG 1988
QY 1141 AGATCCGTACTCCGCGCGGAGGACCTGAGGAGTCGCGATCGACCGGATCGGAACACC 1200
Db |||||
1989 AGATCCGTACTCCGCGCGGAGGACCTGAGGAGTCGCGATCGACCGGATCGGAACACC 2048
QY 1201 TCTCAGAAAGCGCTTAAACAGTCACAGTCGCAAGATC 1239
Db |||||
2049 TCTCAGAAAGCGCTTAAACAGTCACAGTCGCAAGATC 2087


```
A83180
LOCUS A83180 10332 bp DNA linear PAT 21-JAN-2000
DEFINITION Sequence 13 from Patent WO9851788.
ACCESSION A83180
VERSION A83180.1 GI:6732627
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Verheijen, J.H. and Quax, P.H.
TITLE METHOD AND CONSTRUCT FOR INHIBITION OF CELL MIGRATION
JOURNAL Patent: WO 9851788-A 13 19-NOV-1998;
VERHEIJEN JOHAN HENDRIKUS (NL); TNO (NL)
FEATURES
source
1. .10332
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 91.5%; Score 1135; DB 6; Length 10332;
Best Local Similarity 100.0%; Pred. No. 8.6e-217; Indels 0; Gaps 0;
Matches 1135; Conservative 0; Mismatches 0;
QY 6 CACTCTCTCCGCATCGCTGCTCGGAGGCGCAGCTGTTGGGGTGAGTACTCCCTCTGAA 65
DB 4468 CACTCTCTCCGCATCGCTGCTCGGAGGCGCAGCTGTTGGGGTGAGTACTCCCTCTGAA 4527
QY 66 AAGCGGGCATGACTTCTCGGCTAAGATTGTCAGTTTCCAAAACGAGGAGGATTTGATAT 125
DB 4528 AAGCGGGCATGACTTCTCGGCTAAGATTGTCAGTTTCCAAAACGAGGAGGATTTGATAT 4587
QY 126 TCACCTGCGCGCGGTGATGCTTTGAGGTGCGCGCATCCATCTGTCAGAAAAGACAA 185
DB 4588 TCACCTGCGCGCGGTGATGCTTTGAGGTGCGCGCATCCATCTGTCAGAAAAGACAA 4647
QY 186 TCTTTTGTGTTCAAGCTTGTGGCAACGACCCGTAGAGGCGGTGGACAGCAACTTGG 245
DB 4648 TCTTTTGTGTTCAAGCTTGTGGCAACGACCCGTAGAGGCGGTGGACAGCAACTTGG 4707
QY 246 CGATGAGCGCAGGCTTGTGTTTTGTGCGCATGCGCGCTCTCTGCGCGGATGTTA 305
DB 4708 CGATGAGCGCAGGCTTGTGTTTTGTGCGCATGCGCGCTCTCTGCGCGGATGTTA 4767
QY 306 GCTGCACGTATTTCGCGCGCAACGACCGCATTCGGGAAGACGCTGTCGCTCGG 365
DB 4768 GCTGCACGTATTTCGCGCGCAACGACCGCATTCGGGAAGACGCTGTCGCTCGG 4827
QY 366 GCACGAGGTGACGCGCCAAACCGCGTGTGTCAGGTGACAAAGGTCAACGCTGGTGGCTA 425
DB 4828 GCACGAGGTGACGCGCCAAACCGCGTGTGTCAGGTGACAAAGGTCAACGCTGGTGGCTA 4887
QY 426 CTTCTCCGCGTAGGCGCTGTTGGTCCAGCAGAGCGCGCCCTTGGCGCGAGCAATG 485
DB 4888 CTTCTCCGCGTAGGCGCTGTTGGTCCAGCAGAGCGCGCCCTTGGCGCGAGCAATG 4947
QY 486 GCGGTAGGGGTCTAGCTGCTGCTGTCGCGGGGGTCTGCTCCAGGTAAAGACCCCGG 545
DB 4948 GCGGTAGGGGTCTAGCTGCTGCTGTCGCGGGGGTCTGCTCCAGGTAAAGACCCCGG 5007
QY 546 GCAGCAGGCGCGCTCGAAGTAGTCTATCTTCATCTTGGCAAGTCTAGCGCTCGTCC 605
DB 5008 GCAGCAGGCGCGCTCGAAGTAGTCTATCTTCATCTTGGCAAGTCTAGCGCTCGTCC 5067
QY 606 ATGCGCGGCGGCAAGCGCGCTGATGAGTGTGAGTGGGGGACCCCATGTCATGGGCT 665
DB 5068 ATGCGCGGCGGCAAGCGCGCTGATGAGTGTGAGTGGGGGACCCCATGTCATGGGCT 5127
QY 666 GGGTGAGCGCGGAGGCTACATGCGCGCAAAATGTCGTAAACGTAGAGGGGCTCTCTGAT 725
DB 5128 GGGTGAGCGCGGAGGCTACATGCGCGCAAAATGTCGTAAACGTAGAGGGGCTCTCTGAT 5187
```

```
QY 726 TTCCAAGATATGTAGGTPAGCATTTTCCACCGCGATGCTCGCGCGACGTAATCGTATA 785
DB 5188 TTCCAAGATATGTAGGTPAGCATTTTCCACCGCGATGCTCGCGCGACGTAATCGTATA 5247
QY 786 GTTCGTGCGAGGAGCGAGGAGTCCGAGCGAGGTTGCTACGGCGGCTGCTCTGCTC 845
DB 5248 GTTCGTGCGAGGAGCGAGGAGTCCGAGCGAGGTTGCTACGGCGGCTGCTCTGCTC 5307
QY 846 GGAAGACTATCTGCTGAAGATGGCATGTGAGTTGGATGATATGTTGGACGCTGGAAGA 905
DB 5308 GGAAGACTATCTGCTGAAGATGGCATGTGAGTTGGATGATATGTTGGACGCTGGAAGA 5367
QY 906 CGTTGAAGCTGGCGTCTGTGAGACCTACCGCGGTACGCGACGAGAGCGGTAGAGTGC 965
DB 5368 CGTTGAAGCTGGCGTCTGTGAGACCTACCGCGGTACGCGACGAGAGCGGTAGAGTGC 5427
QY 966 GCAGCTTGTGACCAAGCTCGGCGGTGACCTCGAGTCTAGGCGCGAGTCCAGGGTTT 1025
DB 5428 GCAGCTTGTGACCAAGCTCGGCGGTGACCTCGAGTCTAGGCGCGAGTCCAGGGTTT 5487
QY 1026 CTTTGATGATGTCATATCTTCTGCTCCTTTTTCACAGCTCGCGGTAGGACAA 1085
DB 5488 CTTTGATGATGTCATATCTTCTGCTCCTTTTTCACAGCTCGCGGTAGGACAA 5547
QY 1086 ACTCTTCGCGTCTTTCAGTACTCTTGATGCGAAACCCCTCGCGCTCCGAAACG 1140
DB 5548 ACTCTTCGCGTCTTTCAGTACTCTTGATGCGAAACCCCTCGCGCTCCGAAACG 5602
RESULT 8
BD082846
LOCUS BD082846 10332 bp DNA linear PAT 27-AUG-2002
DEFINITION Method and construct for inhibition of cell migration.
ACCESSION BD082846
VERSION BD082846.1 GI:22628456
KEYWORDS JP 2001525669-A/13.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 10332)
AUTHORS Quax, P.H.A. and Verheijen, J.H.
TITLE Method and construct for inhibition of cell migration
JOURNAL Patent: JP 2001525669-A 13 11-DEC-2001;
NEDERLANDSE ORGANISATIE VOOR TOEGEPAST NATUURWETENSCHAPPELIJK
ONDERZOEK TNO
COMMENT
PN JP 2001525669-A/13
PD 11-DEC-2001
PF 11-MAY-1998 JP 1998549077
PR 12-MAY-1997 EP 97201423.7
PI PAULUS HUBERTUS ANDREAS QUAX, JOHAN HENDRIKUS VERHEIJEN PC
C12N9/72, C12N15/62, C07K14/81//C07K19/00
CC Strandedness: Unknown;
CC Topology: Unknown;
FH Key Location/Qualifiers.
FEATURES
source
1. .10332
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 91.5%; Score 1135; DB 6; Length 10332;
Best Local Similarity 100.0%; Pred. No. 8.6e-217; Indels 0; Gaps 0;
Matches 1135; Conservative 0; Mismatches 0;
QY 6 CACTCTCTCCGCATCGCTGCTCGGAGGCGCAGCTGTTGGGGTGAGTACTCCCTCTGAA 65
DB 4468 CACTCTCTCCGCATCGCTGCTCGGAGGCGCAGCTGTTGGGGTGAGTACTCCCTCTGAA 4527
QY 66 AAGCGGGCATGACTTCTCGGCTAAGATTGTCAGTTTCCAAAACGAGGAGGATTTGATAT 125
DB 4528 AAGCGGGCATGACTTCTCGGCTAAGATTGTCAGTTTCCAAAACGAGGAGGATTTGATAT 4587
```

Qy	126	TCAC	TGCGCCGCGGTGATGCCTTTT	TGAGGGTGGCCGATCCATCTG	CTAGAAAAGACAA	185
Db	4588	TCAC	TGCGCCGCGGTGATGCCTTTT	TGAGGGTGGCCGATCCATCTG	CTAGAAAAGACAA	4647
Qy	186	TC	TTTTTTTGTCAAGCTTCGTG	CAACACGACCCGTAGAGGGCTTGC	ACAGCAACTTGG	245
Db	4648	TC	TTTTTTTGTCAAGCTTCGTG	CAACACGACCCGTAGAGGGCTTGC	ACAGCAACTTGG	4707
Qy	246	CGAT	TGAGCGCAGGGTTTGGTTTTT	TGTCG	GATCGGCGCTCTCTTGGCCGCGATGTTTA	305
Db	4708	CGAT	TGAGCGCAGGGTTTGGTTTTT	TGTCG	GATCGGCGCTCTCTTGGCCGCGATGTTTA	4767
Qy	306	GCTG	CAAGTATTCGCGCGCAACG	CACCGCCATTCGGGAAAGACGGTGGT	CGCTCGTCGG	365
Db	4768	GCTG	CAAGTATTCGCGCGCAACG	CACCGCCATTCGGGAAAGACGGTGGT	CGCTCGTCGG	4827
Qy	366	GCAC	CAGGTGCAACGCGGCTTGTG	CAGGGTGACAAGGTCAACCTCGTGGCTA	425	
Db	4828	GCAC	CAGGTGCAACGCGGCTTGTG	CAGGGTGACAAGGTCAACCTCGTGGCTA	4887	
Qy	426	CCTCT	CCGCGTGAAGCGCTCGTTGGT	TCAGCAGAGCGCGCCCTTTCGCGAGCAGAAATG	485	
Db	4888	CCTCT	CCGCGTGAAGCGCTCGTTGGT	TCAGCAGAGCGCGCGCCCTTTCGCGAGCAGAAATG	4947	
Qy	486	CGCGT	TAGGGGTCTAGCTGGCTCTCGT	CCGGGGGTCTGCGTCCACGGTAAAGACCCCGG	545	
Db	4948	CGCGT	TAGGGGTCTAGCTGGCTCTCGT	CCGGGGGTCTGCGTCCACGGTAAAGACCCCGG	5007	
Qy	546	GCAC	CAGCGCGCGCTCGAAGTAGTCTATCTT	TGCATCCTTGCAGGTCTAGCGCCTGCTGCC	605	
Db	5008	GCAC	CAGCGCGCGCTCGAAGTAGTCTATCTT	TGCATCCTTGCAGGTCTAGCGCCTGCTGCC	5067	
Qy	606	ATGCG	CGGCGGCAAGCGCGCTCGTATAGG	TTGAGTGGGGACCCCATGGCATGGGGT	665	
Db	5068	ATGCG	CGGCGGCAAGCGCGCTCGTATAGG	TTGAGTGGGGACCCCATGGCATGGGGT	5127	
Qy	666	GGGT	GAGCGGAGCGGTACATGCGG	CAAAATGTCGTAAACGTAGAGGGGCTCTCTGAGTA	725	
Db	5128	GGGT	GAGCGGAGCGGTACATGCGG	CAAAATGTCGTAAACGTAGAGGGGCTCTCTGAGTA	5187	
Qy	726	TTCCA	AGATATGTAGGGTAGCATCTTCC	ACCGCGATGCTGGCGCGACCGTAATCGTATA	785	
Db	5188	TTCCA	AGATATGTAGGGTAGCATCTTCC	ACCGCGATGCTGGCGCGACCGTAATCGTATA	5247	
Qy	786	GTTCTG	TCGAGGAGCGAGAGGTCTGGGA	CCGAGGTTCTTACGGGCGGGGCTGCTCTGCTC	845	
Db	5248	GTTCTG	TCGAGGAGCGAGAGGTCTGGGA	CCGAGGTTCTTACGGGCGGGGCTGCTCTGCTC	5307	
Qy	846	GCAAC	ACTATCTGCTCAAGATGCATGTG	ATGTTGATGATATCGTTTGGACGCTGGAAGA	905	
Db	5308	GGAAGA	CTATCTGCTCAAGATGCATGTG	ATGTTGATGATATCGTTTGGACGCTGGAAGA	5367	
Qy	906	CGTTG	AACTGGCGCTCTGTGAGACCTT	ACCGCGTCAACGACGAGGAGCGGTAGGAGTCGC	965	
Db	5368	CGTTG	AACTGGCGCTCTGTGAGACCTT	ACCGCGTCAACGACGAGGAGCGGTAGGAGTCGC	5427	
Qy	966	GCAG	CTTGTGACCAAGCTCGCGGTG	ACCTGCACTTAGGGCGCAGTCTCAGGGGTTT	1025	
Db	5428	GCAG	CTTGTGTGACCAAGCTCGCGGTG	ACCTGCACTTAGGGCGCAGTCTCAGGGGTTT	5487	
Qy	1026	CC	TTTGATGATGATCATATTA	CTCTGTCCCTTTTTTTTCCACAGCTCCGGTGTAGGACAA	1085	
Db	5488	CC	TTTGATGATGATCATATTA	CTCTGTCCCTTTTTTTTCCACAGCTCCGGTGTAGGACAA	5547	
Qy	1086	ACTCTT	CGCGGTCTTTTCAGTACTCTT	TGGATCGAAACCCGTCGGGCTCCGAAAG	1140	
Db	5548	ACTCTT	CGCGGTCTTTTCAGTACTCTT	TGGATCGAAACCCGTCGGGCTCCGAAAG	5602	

RESULT 9
AD5001
LOCUS AD5001 11570 bp DNA linear VRL 09-SEP-2004
DEFINITION Adenovirus type 5 left 32% of the genome (coordinates 0% to 32.39%)

ACCESSION	as measured by <ad2>).
VERSION	X02996 J01967 J01968 J01970 J01971 J01972 J01974 J01976 J01977
KEYWORDS	J01978 J01979 K00515 V00025 V00026 V00027 V00029 X02996.1 GI:58484 alternate splicing; DNA polymerase; overlapping genes; polymerase; RNA polymerase III; terminal protein; terminal repeat; transfer RNA.
SOURCE	Human adenovirus type 5
ORGANISM	Human adenovirus type 5
REFERENCE	Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
AUTHORS	1 (bases 1 to 194)
TITLE	Steenbergh, P.H., Maat, J., van Ormondt, H. and Sussenbach, J.S.
JOURNAL	The nucleotide sequence at the termini of adenovirus type 5 DNA
MEDLINE	Nucleic Acids Res. 4 (12), 4371-4389 (1977)
PUBMED	78093872
REFERENCE	600799
AUTHORS	2 (bases 1 to 1574)
TITLE	Van Ormondt, H., Maat, J., De Waard, A. and Van der Eb, A.J.
JOURNAL	The nucleotide sequence of the transforming HpaI-E fragment of
MEDLINE	adenovirus type 5 DNA
PUBMED	Gene 4 (4), 309-328 (1978)
REFERENCE	79128735
AUTHORS	3 (bases 1575 to 2809)
TITLE	Maat, J. and Van Ormondt, H.
JOURNAL	The nucleotide sequence of the transforming HindIII-G fragment of
MEDLINE	adenovirus type 5 DNA. The region between map positions 4.5 (HpaI
PUBMED	site) and 8.0 (HindIII site)
REFERENCE	Gene 6 (1), 75-90 (1979)
AUTHORS	80004833
TITLE	Structure of two spliced mRNAs from the transforming region of
JOURNAL	human subgroup C adenoviruses
MEDLINE	Nature 281 (5733), 694-696 (1979)
PUBMED	81012104
REFERENCE	551290
AUTHORS	5 (bases 10524 to 10696)
TITLE	Thimmappaya, B., Jones, N. and Shenk, T.
JOURNAL	A mutation which alters initiation of transcription by RNA
MEDLINE	polymerase III on the Ad5 chromosome
PUBMED	Cell 18 (4), 947-954 (1979)
REFERENCE	80090080
AUTHORS	51973
TITLE	6 (bases 1 to 6246)
JOURNAL	Broker, T.R.
MEDLINE	Appendix d: nucleotide sequences, transcription and translation
PUBMED	analyses, and restriction endonuclease cleavage maps of group-C
REFERENCE	human adenoviruses
AUTHORS	(in) Toozé, J. (Ed.)
TITLE	DNA TUMOR VIRUSES: 937-1002;
JOURNAL	Cold Spring Harbor Laboratory (1980)
MEDLINE	7 (bases 2804 to 4125)
PUBMED	Maat, J., van Beveren, C.P. and van Ormondt, H.
REFERENCE	The nucleotide sequence of adenovirus type 5 early region E1: the
AUTHORS	region between map positions 8.0 (HindIII site) and 11.8 (SmaI
TITLE	site)
JOURNAL	Gene 10 (1), 27-38 (1980)
MEDLINE	81005097
PUBMED	6250944
REFERENCE	8
AUTHORS	Perriacaudet, M., Le Moullec, J.M. and Pettersson, U.
TITLE	Predicted structure of two adenovirus tumor antigens
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 77 (7), 3778-3782 (1980)
MEDLINE	81054654
PUBMED	6253988
REFERENCE	9 (bases 1 to 4125)
AUTHORS	van Ormondt, H., Maat, J. and van Beveren, C.P.
TITLE	The nucleotide sequence of the transforming early region E1 of
JOURNAL	adenovirus type 5 DNA
MEDLINE	Gene 11 (3-4), 299-309 (1980)
PUBMED	81165537


```
QY 246 CGATGAGCGCAGGGTTGGTTTTTGTGCGCATCGCGCGCTCCTTGGCCGCGATGTTTA 305
DB 6288 CGATGAGCGCAGGGTTGGTTTTTGTGCGCATCGCGCGCTCCTTGGCCGCGATGTTTA 6347
QY 306 GCTGCACGTATTTCGCGCGCAACCGCACCGCATTCGGGAAAGACGGTGGTGGCTCGTGG 365
DB 6348 GCTGCACGTATTTCGCGCGCAACCGCACCGCATTCGGGAAAGACGGTGGTGGCTCGTGG 6407
QY 366 GCACAGGTGCACGCGCCAAACCGCGGTGTGACAGGTGACAAAGTCAACGCTGTGGGCTA 425
DB 6408 GCACAGGTGCACGCGCCAAACCGCGGTGTGACAGGTGACAAAGTCAACGCTGTGGGCTA 6467
QY 426 CCTCTCCGCTAGGCGCTGTTGGTTCACAGAGGCGCGCCCTTGGCGGACGAGATG 485
DB 6468 CCTCTCCGCTAGGCGCTGTTGGTTCACAGAGGCGCGCCCTTGGCGGACGAGATG 6527
QY 486 GCGGTAGGGGGTCTAGCTCGCTCTCGTCCGGGGGTCTGCGTCCACCGGTAAAGACCCCGG 545
DB 6528 GCGGTAGGGGGTCTAGCTCGCTCTCGTCCGGGGGTCTGCGTCCACCGGTAAAGACCCCGG 6587
QY 546 GCAGCAGCGCGCTCGAAGTAGTCTATCTTGCAATCTTGGCAATCTAGCGCTGCTGCC 605
DB 6588 GCAGCAGCGCGCTCGAAGTAGTCTATCTTGCAATCTTGGCAATCTAGCGCTGCTGCC 6647
QY 606 ATGCGCGCGCGCAAGCGCGCTCGTATGGTTCAGTGGGGACCCCATGGCATGGGGT 665
DB 6648 ATGCGCGCGCGCAAGCGCGCTCGTATGGTTCAGTGGGGACCCCATGGCATGGGGT 6707
QY 666 GGGTGAGCGCGAGCGGTACATCGCGCAAAATGTCGTAAACGTAGAGGGGTCTCTGAGTA 725
DB 6708 GGGTGAGCGCGAGCGGTACATCGCGCAAAATGTCGTAAACGTAGAGGGGTCTCTGAGTA 6767
QY 726 TTCCAAGATATAGGGTAGCATTTTCCACCGCGATGCTGCGCGCACGTAATCGTATA 785
DB 6768 TTCCAAGATATAGGGTAGCATTTTCCACCGCGATGCTGCGCGCACGTAATCGTATA 6827
QY 786 GTTCTGCGAGGACGAGGAGGTGCGGACCGAGGTGCTACGGCGGGGTCTCTGCTC 845
DB 6828 GTTCTGCGAGGACGAGGAGGTGCGGACCGAGGTGCTACGGCGGGGTCTCTGCTC 6887
QY 846 GGAAGCATCTCTGCTGAAGATGGCATGTGAGTTCGGATGATATGTTGGACGCTGGAAGA 905
DB 6888 GGAAGCATCTCTGCTGAAGATGGCATGTGAGTTCGGATGATATGTTGGACGCTGGAAGA 6947
QY 906 CTTTGAAGCTGCGCTCTGTGAGACTTACCGGTACGACGACGACGAGAGCGGTAGGAGTGC 965
DB 6948 CTTTGAAGCTGCGCTCTGTGAGACTTACCGGTACGACGACGACGAGAGCGGTAGGAGTGC 7007
QY 966 GCAGCTTTGTTGACCGCTCGCGGTGACCTGACGCTTAGGGCGCAGTAGTCCAGGGTTT 1025
DB 7008 GCAGCTTTGTTGACCGCTCGCGGTGACCTGACGCTTAGGGCGCAGTAGTCCAGGGTTT 7067
QY 1026 CTTTGATGATGTATACCTATCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1085
DB 7068 CTTTGATGATGTATACCTATCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7127
QY 1086 ACTCTTCGCGGTCTTTCAGTACTCTTGGATCGGAAACCCGCTCGGCCTCCGAACG 1140
DB 7128 ACTCTTCGCGGTCTTTCAGTACTCTTGGATCGGAAACCCGCTCGGCCTCCGAACG 7182
```

```
RESULT 10
CQ854904
LOCUS 31976 bp DNA linear PAT 23-AUG-2004
DEFINITION
Sequence 1 from Patent WO2004066947.
ACCESSION
CQ854904
VERSION
CQ854904.1 GI:51510464
KEYWORDS
unidentified adenovirus
SOURCE
unidentified adenovirus
ORGANISM
Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
REFERENCE
1
AUTHORS Hu, F. and Wu, B.
```

```
TITLE Therapy for primary and metastatic cancers
JOURNAL Patent: WO 2004066947-A 1 12-AUG-2004;
Shanghai Sunway Biotech Co Ltd (CN)
FEATURES
source 1. 31976
/organism="unidentified adenovirus"
/mol_type="unassigned DNA"
/db_xref="taxon:10535"
ORIGIN
Query Match 91.5%; Score 1135; DB 6; Length 31976;
Best Local Similarity 100.0%; Pred. No. 7.9e-217;
Matches 1135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 6 CACTCTCTTCGCGCATCGCTGTCTCGAGGGCCAGCTGTTGGGGTGAGTACTCCCTCTGAA 65
DB 5222 CACTCTCTTCGCGCATCGCTGTCTCGAGGGCCAGCTGTTGGGGTGAGTACTCCCTCTGAA 5281
QY 66 AAGCGGGCATGACTTCTCGCTAAGATTGTCAGTTTCCAAAACGAGAGGATTTGATAT 125
DB 5282 AAGCGGGCATGACTTCTCGCTAAGATTGTCAGTTTCCAAAACGAGAGGATTTGATAT 5341
QY 126 TCACCTCGCCCGCGCTGATGCTTTGAGGGTGGCGCATCCATCTGTGCAGAAAGACAA 185
DB 5342 TCACCTCGCCCGCGCTGATGCTTTGAGGGTGGCGCATCCATCTGTGCAGAAAGACAA 5401
QY 186 TCTTTTGTGTCTAAGCTTGGTGGCAACGACCCGTAGAGGGCTTGGACAGCAACTTGG 245
DB 5402 TCTTTTGTGTCTAAGCTTGGTGGCAACGACCCGTAGAGGGCTTGGACAGCAACTTGG 5461
QY 246 CGATGGAGCGCAGGGTTGGTTTTTGTGCGCATCGCGCGCTCCTTGGCCGCGATGTTA 305
DB 5462 CGATGGAGCGCAGGGTTGGTTTTTGTGCGCATCGCGCGCTCCTTGGCCGCGATGTTA 5521
QY 306 GCTCAGCATATTCCGCGCGCAACGACCGCATTCGGAAGACCGGTGCGCTCGTCGG 365
DB 5522 GCTCAGCATATTCCGCGCGCAACGACCGCATTCGGAAGACCGGTGCGCTCGTCGG 5581
QY 366 GCACAGGTGCACGCGCCAAACCGCGTGTGTCAGGGTGACAAAGTCAACGCTGTGGCTA 425
DB 5582 GCACAGGTGCACGCGCCAAACCGCGTGTGTCAGGGTGACAAAGTCAACGCTGTGGCTA 5641
QY 426 CTTCTCCGCTAGCGCTCGTGTGTCAGCAGAGCGCGCCCTTGGCGGACGAGATG 485
DB 5642 CTTCTCCGCTAGCGCTCGTGTGTCAGCAGAGCGCGCCCTTGGCGGACGAGATG 5701
QY 486 CCGGTAGGGGTCTAGCTGCGTCTGCTCCGGGGGTCTGCGTCCACGCTAAAGACCCCGG 545
DB 5702 CCGGTAGGGGTCTAGCTGCGTCTGCTCCGGGGGTCTGCGTCCACGCTAAAGACCCCGG 5761
QY 546 GCAGCAGCGCGCTCGAAGTGTATCTTGTGCAATCTTGCAGTCTAGCGCTGTGCTCC 605
DB 5762 GCAGCAGCGCGCTCGAAGTGTATCTTGTGCAATCTTGCAGTCTAGCGCTGTGCTCC 5821
QY 606 ATGCGCGCGCGCAAGCGCGCTCGTATGGTTCAGTGGGGACCCCATGGCATGGGGT 665
DB 5822 ATGCGCGCGCGCAAGCGCGCTCGTATGGTTCAGTGGGGACCCCATGGCATGGGGT 5881
QY 666 GGGTGAGCGGAGCGGTACATGCGCAAAATGTCGTAACGATAGAGGGGTCTCTGAGTA 725
DB 5882 GGGTGAGCGGAGCGGTACATGCGCAAAATGTCGTAACGATAGAGGGGTCTCTGAGTA 5941
QY 726 TTCCAAGATATGATAGGGTAGCATCTTCCACCGCGATGCTGGCGCGCACGCTAAATCGTATA 785
DB 5942 TTCCAAGATATGATAGGGTAGCATCTTCCACCGCGATGCTGGCGCGCACGCTAAATCGTATA 6001
QY 786 GTTCTGCGAGGACGAGGAGGTTCGGACCGAGTTGCTACGCGCGCGCTCTGCTC 845
DB 6002 GTTCTGCGAGGACGAGGAGGTTCGGACCGAGTTGCTACGCGCGCGCTCTGCTC 6061
QY 846 GGAAGATCTATCTGCTGAAAGTGGCATGTGAGTTGGATGATATGTTGGACGCTGGAAGA 905
DB 6062 GGAAGATCTATCTGCTGAAAGTGGCATGTGAGTTGGATGATATGTTGGACGCTGGAAGA 6121
```

5642	Db	CCTCTCCGCGTAGCGCTCGTTGGTTCAGCAGAGGCGCGCCCTCTTGC	CGCAGCAATG	5701
486	Qy	GCCTTAGGGGGTCTAGCTGCGTCTCGTCCGGGGGGTCTGCGTCCACGGTAAAGACCCCGG	545	
5702	Db	CGCGTAGGGGGTCTAGCTGCGTCTCGTCCGGGGGGTCTGCGTCCACGGTAAAGACCCCGG	5761	
546	Qy	GCAGCAGCGCGCGCTGAAAGTAGTCTATCTTGCATCTTGCAGAGTCTAGCGCTGCTGCC	605	
5762	Db	GCACGAGCGCGCGCTGAAAGTAGTCTATCTTGCATCTTGCAGAGTCTAGCGCTGCTGCC	5821	
606	Qy	ATCGCGGGCGCGCAAGCGCGCGCTCGTATGGGTTCAGTGGGGACCCACCGCATGGCGTGGGT	665	
5822	Db	ATCGCGGGCGCGCAAGCGCGCGCTCGTATGGGTTCAGTGGGGACCCACCGCATGGCGTGGGT	5881	
666	Qy	GGGTGAGCGCGGAGCGGTATCATGCGCAAAATGCTGTAACAGTAGAGGGGCTCTCTGAGTA	725	
5882	Db	GGGTGAGCGCGGAGCGGTATCATGCGCAAAATGCTGTAACAGTAGAGGGGCTCTCTGAGTA	5941	
726	Qy	TTTCAAGATATGTAGGGTAGCATCTTCCACCGCGGATCTGGCGCGCACGTAATCGTATA	785	
5942	Db	TTTCAAGATATGTAGGGTAGCATCTTCCACCGCGGATCTGGCGCGCACGTAATCGTATA	6001	
786	Qy	GTTCGTGCGAGGAGCGAGGAGTCCGGACCGAGGTTCCTACGGCGGGCTGCTCTGCTC	845	
6002	Db	GTTCGTGCGAGGAGCGAGGAGTCCGGACCGAGGTTCCTACGGCGGGCTGCTCTGCTC	6061	
846	Qy	GGAAGACTATCTGCTGGAAGATGGCATGTGAGTTGCATGATATGTTGGAACGCTGGAAGA	905	
6062	Db	GGAAGACTATCTGCTGGAAGATGGCATGTGAGTTGCATGATATGTTGGAACGCTGGAAGA	6121	
906	Qy	CGTTGAAGCTGGCGTCTGTGACACCTACCGCTCACGCACGACGAGGAGCGTAGGAGTCGC	965	
6122	Db	CGTTGAAGCTGGCGTCTGTGACACCTACCGCTCACGCACGACGAGGAGCGTAGGAGTCGC	6181	
966	Qy	GCAGCTTGTGTGACCAAGCTCGCGGTCACCTGCAAGCTCTAGGGCGCAGTAGTCAGAGGTTT	1025	
6182	Db	GCAGCTTGTGTGACCAAGCTCGCGGTCACCTGCAAGCTCTAGGGCGCAGTAGTCAGAGGTTT	6241	
1026	Qy	CCGTGATGATGTCATACCTATCTGTCCTCTTTTTCACAGCTCGCGGTGAGACAA	1085	
6242	Db	CCGTGATGATGTCATACCTATCTGTCCTCTTTTTCACAGCTCGCGGTGAGACAA	6301	
1086	Qy	ACTCTTCGCGGCTTTTCCAGTACTCTTGGATCGAAACCCGTCGGCTCCGAACG	1140	
6302	Db	ACTCTTCGCGGCTTTTCCAGTACTCTTGGATCGAAACCCGTCGGCTCCGAACG	6356	
RESULT 12				
BD268216				
LOCUS		32480 bp	DNA	linear
DEFINITION		Adenovirus vector, packaging cell line, composition and method for production and use.		PAT 17-JUL-2003
ACCESSION		BD268216		
VERSION		BD268216.1		GI:33077984
KEYWORDS		JP 2002534130-A/20.		
SOURCE		unidentified adenovirus		
ORGANISM		unidentified adenovirus		
REFERENCE		Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus. 1 (bases 1 to 32480)		
AUTHORS		Nemerow,G.R., Seggern,D.J.V., Hallenbeck,P.L., Stevenson,S.C. and Skripchenko, Y.		
TITLE		Adenovirus vector, packaging cell line, composition and method for production and use		
JOURNAL		Patent: JP 2002534130-A 20 15-OCT-2002;		
COMMENT		NOVARTIS AG,THE SCRIPPS RESEARCH INSTITUTE		
		OS Adenovirus		
		PN JP 2002534130-A/20		
		PD 15-OCT-2002		
		PF 14-JAN-2000		JP 2000593765
		PR 14-JAN-1999		US 60/115920
		PI GLEN ROBERT NEMEROW,DANIEL J VON SEGGERN,PAUL L HALLENBECK, PI SUSAN C STEVENSON,YELENA SKRIPCENKO		

PC C12N15/09,A61K35/76,A61K48/00,A61P35/00,A61P43/00,A61P43/00, 845
PC C12N5/10, 845
PC C12N7/00,C12Q1/68,G01N33/53,G01N33/566,C12N15/00,C12N5/00 CC 8031
Adenovirus vector, packaging cell line, 8031
composition and method 905
CC production and use 8091
CC key Location/Qualifiers 8091
FT source 1.32480 965
FT Location/Qualifiers /organism='Adenovirus', 8151
FEATURES 1.32480 1025
source /organism='unidentified adenovirus' 8211
/mol_type='genomic DNA' 1085
/db_xref='taxon:10535' 8271

ORIGIN 8271

Query Match 91.5%; Score 1135; DB 6; Length 32480; 1085
Best Local Similarity 100.0%; Pred. No. 7.9e-217; 8271
Matches 1135; Conservative 0; Mismatches 0; Indels 0; Gaps 0; 1085

QY 6 CACTCTCTTCGCGATCGCTGCTGCGAGGCGCAGCTGTTGGGGTGAGTACTCCCTCTGAA 65
DB CACTCTCTTCGCGATCGCTGCTGCGAGGCGCAGCTGTTGGGGTGAGTACTCCCTCTGAA 7251

QY 66 AAGCGGCGATGACTCTTCGCGTAAAGATTGTGATTTCCAAAACAGAGGAGATTGATAT 125
DB AAGCGGCGATGACTCTTCGCGTAAAGATTGTGATTTCCAAAACAGAGGAGATTGATAT 7311

QY 126 TCACCTGCGCGCGGCGTGTGAGGTTGGCGGCGATCCATCTGTCAGAAAGACAA 185
DB TCACCTGCGCGCGGCGTGTGAGGTTGGCGGCGATCCATCTGTCAGAAAGACAA 7371

QY 186 TCTTTTGTGTTCAAGCTTGTGCGCAACCGCTAGAGGCGGTTGGACAGCAACTTGG 245
DB TCTTTTGTGTTCAAGCTTGTGCGCAACCGCTAGAGGCGGTTGGACAGCAACTTGG 7431

QY 246 CGATGAGCGAGGCGTGTGTTGTCGCGATCGCGCGCTCTCTTGGCGCGATGTTTA 305
DB CGATGAGCGAGGCGTGTGTTGTCGCGATCGCGCGCTCTCTTGGCGCGATGTTTA 7491

QY 306 GCTGACGATTTCGCGCGCAACCGCTGTTGGGAAAGACGCTGTCGCTCGG 365
DB GCTGACGATTTCGCGCGCAACCGCTGTTGGGAAAGACGCTGTCGCTCGG 7551

QY 366 GCACGAGTGCAACCGCGCTGTTGTCAGGTTGACAAAGTCAACGCTGTTGGCTA 425
DB GCACGAGTGCAACCGCGCTGTTGTCAGGTTGACAAAGTCAACGCTGTTGGCTA 7611

QY 426 CCTCTCGCGTAGGCGCTGTTGTCAGAGAGCGCGCGCTTGGCGAGCAAGATG 485
DB CCTCTCGCGTAGGCGCTGTTGTCAGAGAGCGCGCGCTTGGCGAGCAAGATG 7671

QY 486 GCGGTAGGGGCTAGCTGCTGTCGCGGGGCTGCGTCCACGCTGAAGACCCCGG 545
DB GCGGTAGGGGCTAGCTGCTGTCGCGGGGCTGCGTCCACGCTGAAGACCCCGG 7731

QY 546 GCAGCAGGCGCGCTGCAAGTAGTCTATCTTCATCTTGGCAAGTCTAGCGCTGCTGCC 605
DB GCAGCAGGCGCGCTGCAAGTAGTCTATCTTCATCTTGGCAAGTCTAGCGCTGCTGCC 7791

QY 606 ATGCGGGCGCGCAACGCGCGCTGATGAGTTGAGTGGGGAACCCCATGCGATGGGGT 665
DB ATGCGGGCGCGCAACGCGCGCTGATGAGTTGAGTGGGGAACCCCATGCGATGGGGT 7851

QY 666 GGGTCAGCGCGAGCGCTACATCGCGAAATGTCGTAAACGTAGAGGGGCTCTCTGAGTA 725
DB GGGTCAGCGCGAGCGCTACATCGCGAAATGTCGTAAACGTAGAGGGGCTCTCTGAGTA 7911

QY 726 TTCCAAAGATATAGGGTAGCATCTTCCACCGCGGATGTCGCGCGCAAGTAAATCGTATA 785
DB TTCCAAAGATATAGGGTAGCATCTTCCACCGCGGATGTCGCGCGCAAGTAAATCGTATA 7971

QY 786 GTTCTGTCGAGGAGCGAGGAGGCTCGGACCGAGTTGCTTACGCGCGGCTGCTCTGCTC 845
DB GTTCTGTCGAGGAGCGAGGAGGCTCGGACCGAGTTGCTTACGCGCGGCTGCTCTGCTC 8031

QY 846 GGAAGACTATCTGCTGAAAGATGCGATGTGAGTTGGATGATATGTTGGACGCTTGAAGA 905
DB GGAAGACTATCTGCTGAAAGATGCGATGTGAGTTGGATGATATGTTGGACGCTTGAAGA 8091

QY 906 CGTTGAAGCTGGCGTCTGTGAGACCTACCGCGTCACGACGAAAGAGCGTAGAGTCGC 965
DB CGTTGAAGCTGGCGTCTGTGAGACCTACCGCGTCACGACGAAAGAGCGTAGAGTCGC 8151

QY 966 GCACCTTGTGACCGAGCTCGCGGCTGACCTGACGCTTAGGGCGCAGTAGTCCAGGGTTT 1025
DB GCACCTTGTGACCGAGCTCGCGGCTGACCTGACGCTTAGGGCGCAGTAGTCCAGGGTTT 8211

QY 1026 CCTTGATGATGTATATCTTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1085
DB CCTTGATGATGTATATCTTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 8271

QY 1086 ACTCTTCGCGGCTCTTCCAGTACTCTTGGATCGGAAACCGCTCGGCTCGGAAACG 1140
DB ACTCTTCGCGGCTCTTCCAGTACTCTTGGATCGGAAACCGCTCGGCTCGGAAACG 8326

RESULT 13
AR343138 32798 bp DNA linear PAT 17-AUG-2003
LOCUS Sequence 1 from patent US 6579522.
DEFINITION AR343138
ACCESSION AR343138
VERSION AR343138.1 GI:33738640
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 32798)
AUTHORS Brought,D.E., King,C.R., Kovesdi,I. and Schaible,J.J.
TITLE Replication deficient adenoviral TNF vector
JOURNAL Patent: US 6579522-A 1 17-JUN-2003;
FEATURES Location/Qualifiers
source 1.32798
/organism='unknown'
/mol_type='genomic DNA'

ORIGIN

Query Match 91.5%; Score 1135; DB 6; Length 32798;
Best Local Similarity 100.0%; Pred. No. 7.9e-217;
Matches 1135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 CACTCTCTTCGCGATCGCTGCTGCGAGGCGCAGCTGTTGGGGTGAGTACTCCCTCTGAA 65
DB CACTCTCTTCGCGATCGCTGCTGCGAGGCGCAGCTGTTGGGGTGAGTACTCCCTCTGAA 5114

QY 66 AAGCGGCGATGACTCTTCGCGTAAAGATTGTGATTTCCAAAACAGAGGAGATTGATAT 125
DB AAGCGGCGATGACTCTTCGCGTAAAGATTGTGATTTCCAAAACAGAGGAGATTGATAT 5174

QY 126 TCACCTGCGCGCGGCTGATGCTTTGAGGGTGCGCATCCATCTGTCAGAAAGACAA 185
DB TCACCTGCGCGCGGCTGATGCTTTGAGGGTGCGCATCCATCTGTCAGAAAGACAA 5234

QY 186 TCTTTTGTGTTCAAGCTTGTGTCGCAACCGCTGAGAGGCGTTGACAGCAACTTGG 245
DB TCTTTTGTGTTCAAGCTTGTGTCGCAACCGCTGAGAGGCGTTGACAGCAACTTGG 5294

QY 246 CGATGAGCGCGAGGTTGGTTTGTGCGCATCGGCGCTCTCTGSCCGGATGTTTA 305
DB CGATGAGCGCGAGGTTGGTTTGTGCGCATCGGCGCTCTCTGSCCGGATGTTTA 5354

QY 306 GCTGACGATTTCGCGCGCAACCGCACCGCAATTCGGGAAAGACGCTGTCGCTCGG 365
DB GCTGACGATTTCGCGCGCAACCGCACCGCAATTCGGGAAAGACGCTGTCGCTCGG 5414

QY 366 GCACACAGGTGACGCGCCCAACCGCGGTGTGACAGGGTGAACAAGTCAACGCTGGTGCTA 425
Db 5415 GCACACAGGTGACGCGCCCAACCGCGGTGTGACAGGGTGAACAAGTCAACGCTGGTGCTA 5474
QY 426 CTTCTCCGCTAGGCGCTCGTTGGTCCAGACAGGCGCGCCCTTGGCGGAGCAGAATG 485
Db 5475 CTTCTCCGCTAGGCGCTCGTTGGTCCAGACAGGCGCGCCCTTGGCGGAGCAGAATG 5534
QY 486 GCGTAGGGGGTCTAGCTCGCTCTCGTCCGGGGGTCTGCTCCACGCTAAAGACCCCGG 545
Db 5535 GCGTAGGGGGTCTAGCTCGCTCTCGTCCGGGGGTCTGCTCCACGCTAAAGACCCCGG 5594
QY 546 GCACAGGCGCGGCTCGAAGTAGTCTATCTTGATCTTGCATCTTGCAGCTAGCGCTGCTGCC 605
Db 5595 GCACAGGCGCGGCTCGAAGTAGTCTATCTTGATCTTGCATCTTGCAGCTAGCGCTGCTGCC 5654
QY 606 ATGCGCGGCGGCAAGCGCGCTCGTATGGGTTCAGTGGGGGACCCCATGCGATGGGT 665
Db 5655 ATGCGCGGCGGCAAGCGCGCTCGTATGGGTTCAGTGGGGGACCCCATGCGATGGGT 5714
QY 666 GGGTAGCGCGGAGGCGGTACATCCCGCAATGTCTGTAACGTAGAGGGGCTCTCTGAGTA 725
Db 5715 GGGTAGCGCGGAGGCGGTACATCCCGCAATGTCTGTAACGTAGAGGGGCTCTCTGAGTA 5774
QY 726 TTCCAAGATATGAGGTAGATCTTTCCACCGCGATCTGCGCGCACGTAATCGTATA 785
Db 5775 TTCCAAGATATGAGGTAGATCTTTCCACCGCGATCTGCGCGCACGTAATCGTATA 5834
QY 786 GTTCTGCGAGGAGCGAGGTGCGGACCGAGGTGCTACGCGGGGCTCTCTGCTC 845
Db 5835 GTTCTGCGAGGAGCGAGGTGCGGACCGAGGTGCTACGCGGGGCTCTCTGCTC 5894
QY 846 GGAAGATATCTGCTGAAGATGGCATGTGAGTGGATGATATGTTGACCGCTCGAAGA 905
Db 5895 GGAAGATATCTGCTGAAGATGGCATGTGAGTGGATGATATGTTGACCGCTCGAAGA 5954
QY 906 CTTGAAAGTGTGAGACCTTACCGCTCAGCTACGCAAGAGGCGGTAGGAGTCCG 965
Db 5955 CTTGAAAGTGTGAGACCTTACCGCTCAGCTACGCAAGAGGCGGTAGGAGTCCG 6014
QY 966 GCAGCTTGTGACAGCTCGGCGGTGACCTGACGCTTAGGGCGAGTAGTCCAGGGTTT 1025
Db 6015 GCAGCTTGTGACAGCTCGGCGGTGACCTGACGCTTAGGGCGAGTAGTCCAGGGTTT 6074
QY 1026 CTTGATGATGTCACTATCTCTGCTCCCTTTTTCACAGCTCGCGGTTGAGACAA 1085
Db 6075 CTTGATGATGTCACTATCTCTGCTCCCTTTTTCACAGCTCGCGGTTGAGACAA 6134
QY 1086 ACTCTTCGCGGTCTTTCCAGTACTCTTGGATCGAAACCGCTCGGCTCCGAACG 1140
Db 6135 ACTCTTCGCGGTCTTTCCAGTACTCTTGGATCGAAACCGCTCGGCTCCGAACG 6189

RESULT 14

AX382187 32798 bp DNA linear PAT 18-MAR-2002
LOCUS Sequence 1 from Patent WO200906.

DEFINITION AX382187

ACCESSION AX382187.1 GI:19576990

VERSION Human adenovirus type 5

KEYWORDS Human adenovirus type 5

SOURCE Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.

ORGANISM Brough, D.E., King, C.R. and Kovacs, I.

REFERENCE Replication deficient adenoviral tnF vector

AUTHORS Patent: WO 020906-A 1 03-JAN-2002;

JOURNAL GENVEC, INC. (US)

FEATURES Location/Qualifiers

1..32798

/organism="Human adenovirus type 5"

/mol_type="unassigned DNA"

/db_xref="taxon:28285"

ORIGIN

Query Match 91.5%; Score 1135; DB 6; Length 32798;
Best Local Similarity 100.0%; Pred. No. 7.9e-217;
Matches 1135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 CACTCTCTTCGCGATCGCTGCTGCGAGGCGCAGCTGTTGGGGTGAGTACTCCCTCTGAA 65
Db 5055 CACTCTCTTCGCGATCGCTGCTGCGAGGCGCAGCTGTTGGGGTGAGTACTCCCTCTGAA 5114
QY 66 AAGCGGGCATGACTTCTCGCTAAGATTCTCAGATTTCGAGTTTCCAAAACAGAGGAGTTTGAAT 125
Db 5115 AAGCGGGCATGACTTCTCGCTAAGATTCTCAGATTTCGAGTTTCCAAAACAGAGGAGTTTGAAT 5174
QY 126 TCACCTGCGCGCGGTGATGCTTTCAGGGTGGCGGCATCCCATCTGCTCAGAAAACACAA 185
Db 5175 TCACCTGCGCGCGGTGATGCTTTCAGGGTGGCGGCATCCCATCTGCTCAGAAAACACAA 5234
QY 186 TCTTTTGTGTCAAGCTTGTGTCGCAACCGCTAGAGGCGTTGGACAGCAACTTGG 245
Db 5235 TCTTTTGTGTCAAGCTTGTGTCGCAACCGCTAGAGGCGTTGGACAGCAACTTGG 5294
QY 246 CGATGGAGGCGAGGGTTTGGTTTGTTCGCGATCGCGCGCTCTTGGCGCGATGTTTA 305
Db 5295 CGATGGAGGCGAGGGTTTGGTTTGTTCGCGATCGCGCGCTCTTGGCGCGATGTTTA 5354
QY 306 GCTGCACGTATTTCGCGCGCAACCGCACCGCATTCGGGAAAGACGTTGGTGGCTCGTCGG 365
Db 5355 GCTGCACGTATTTCGCGCGCAACCGCACCGCATTCGGGAAAGACGTTGGTGGCTCGTCGG 5414
QY 366 GCACAGGTGCAACCGCGCTTGTGACGGTTCACAGGTCAACGCTGGTGGCTA 425
Db 5415 GCACAGGTGCAACCGCGCTTGTGACGGTTCACAGGTCAACGCTGGTGGCTA 5474
QY 426 CCTCTCCGCTAGGCGCTGTTGTCACAGAGGCGCGCGCTTGGCGGAGCAAGT 485
Db 5475 CCTCTCCGCTAGGCGCTGTTGTCACAGAGGCGCGCGCTTGGCGGAGCAAGT 5534
QY 486 GGGTAGGGGTCTAGCTCGCTCCGCGGGTCTGCGTCCAGCGTCAACGCTGAGGAGT 545
Db 5535 GGGTAGGGGTCTAGCTCGCTCCGCGGGTCTGCGTCCAGCGTCAACGCTGAGGAGT 5594
QY 546 GCACAGGCGCGCTGCAAGTAGTCTATCTTGCATCTTTCAGGTCTAGCGCTGTCGCC 605
Db 5595 GCACAGGCGCGCTGCAAGTAGTCTATCTTGCATCTTTCAGGTCTAGCGCTGTCGCC 5654
QY 606 ATGCGCGGCGGCAAGCGCGCTCGTATGGGTGAGTGGGGGACCCCATGCGATGGGT 665
Db 5655 ATGCGCGGCGGCAAGCGCGCTCGTATGGGTGAGTGGGGGACCCCATGCGATGGGT 5714
QY 666 GGGTAGGCGGAGGCGGTACATCCCGCAATGTCTGTAACGTAGAGGGGCTCTCTGAGTA 725
Db 5715 GGGTAGGCGGAGGCGGTACATCCCGCAATGTCTGTAACGTAGAGGGGCTCTCTGAGTA 5774
QY 726 TTCCAAGATATGAGGTAGATCTTTCCACCGCGATCTGCGCGCACGTAATCGTATA 785
Db 5775 TTCCAAGATATGAGGTAGATCTTTCCACCGCGATCTGCGCGCACGTAATCGTATA 5834
QY 786 GTTCTGCGAGGAGCGAGGTGCGGACCGAGGTGCTACGCGGGGCTCTCTGCTC 845
Db 5835 GTTCTGCGAGGAGCGAGGTGCGGACCGAGGTGCTACGCGGGGCTCTCTGCTC 5894
QY 846 GGAAGATATCTGCTGAAGATGGCATGTGAGTGGATGATATGTTGACCGCTCGAAGA 905
Db 5895 GGAAGATATCTGCTGAAGATGGCATGTGAGTGGATGATATGTTGACCGCTCGAAGA 5954
QY 906 CTTGAAAGTGTGAGACCTTACCGCTCAGCTACGCAAGAGGCGGTAGGAGTCCG 965
Db 5955 CTTGAAAGTGTGAGACCTTACCGCTCAGCTACGCAAGAGGCGGTAGGAGTCCG 6014
QY 966 GCAGCTTGTGACAGCTCGGCGGTGACCTGACGCTTAGGGCGAGTAGTCCAGGGTTT 1025
Db 6015 GCAGCTTGTGACAGCTCGGCGGTGACCTGACGCTTAGGGCGAGTAGTCCAGGGTTT 6074
QY 1026 CTTGATGATGTCACTATCTCTGCTCCCTTTTTCACAGCTCGCGGTTGAGACAA 1085
Db 6075 CTTGATGATGTCACTATCTCTGCTCCCTTTTTCACAGCTCGCGGTTGAGACAA 6134
QY 1086 ACTCTTCGCGGTCTTTCCAGTACTCTTGGATCGAAACCGCTCGGCTCCGAACG 1140
Db 6135 ACTCTTCGCGGTCTTTCCAGTACTCTTGGATCGAAACCGCTCGGCTCCGAACG 6189

```
QY 1026 CCTTGATGATGTCATCTATCTCTGCTCCCTTTTTCACAGCTCGCGTTGAGGACAA 1085
Db 6075 CCTTGATGATGTCATCTATCTCTGCTCCCTTTTTCACAGCTCGCGTTGAGGACAA 6134
QY 1086 ACTCTTCGCGGTCTTTCCAGTACTCTTGATCGGAAACCGCTCGCCCTCCGAAACG 1140
Db 6135 ACTCTTCGCGGTCTTTCCAGTACTCTTGATCGGAAACCGCTCGCCCTCCGAAACG 6189

RESULT 15
CQ854906 32802 bp DNA linear PAT 23-AUG-2004
DEFINITION Sequence 3 from Patent WO2004066947.
ACCESSION CQ854906
VERSION CQ854906.1 GI:51510466
KEYWORDS
SOURCE
ORGANISM
unidentified adenovirus
unidentified adenovirus
Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
1
REFERENCE
AUTHORS Hu,F. and Wu,B.
TITLE Therapy for primary and metastatic cancers
JOURNAL Patent: WO 2004066947-A 3 12-AUG-2004;
Shanghai Sunway Biotech Co Ltd (CN)
FEATURES
source
1..32802
/organism="unidentified adenovirus"
/mol_type="unassigned DNA"
/db_xref="taxon:10535"

ORIGIN
Query Match 91.5%; Score 1135; DB 6; Length 32802;
Best Local Similarity 100.0%; Pred. No. 7.9e-217;
Matches 1135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 CACTCTTCGCGATCGCTGCTCGAGGCGCAGCTGTTGGGTGAGTACTCCCTCTGAA 65
Db 6048 CACTCTTCGCGATCGCTGCTCGAGGCGCAGCTGTTGGGTGAGTACTCCCTCTGAA 6107

QY 66 AAGCGGGCATGATCTTCGCGTAAGATTGTCAAGTTTCCAAAAACGAGGAGGATTGATAT 125
Db 6108 AAGCGGGCATGATCTTCGCGTAAGATTGTCAAGTTTCCAAAAACGAGGAGGATTGATAT 6167

QY 126 TCACCTGGCCCGCGGTGATGCTTTGAGGTGGCCGATCCATCTGGTCAAGAAAGACAA 185
Db 6168 TCACCTGGCCCGCGGTGATGCTTTGAGGTGGCCGATCCATCTGGTCAAGAAAGACAA 6227

QY 186 TCTTTTGTGTCAAGCTTGTGGCAACGACCCGTAGAGGCGTTGGACAGCAACTTGG 245
Db 6228 TCTTTTGTGTCAAGCTTGTGGCAACGACCCGTAGAGGCGTTGGACAGCAACTTGG 6287

QY 246 CGATGGAGCGGAGGTTGTTGTTTGTGCGATCGCGCGCTCTCTTGGCGCGATGTTTA 305
Db 6288 CGATGGAGCGGAGGTTGTTGTTTGTGCGATCGCGCGCTCTCTTGGCGCGATGTTTA 6347

QY 306 GGTGACGATATTCGCGCAACGACCGCATTCGCGGAAAGACGTTGTTGCGCTCGTCGG 365
Db 6348 GGTGACGATATTCGCGCGCAACGACCGCATTCGCGGAAAGACGTTGTTGCGCTCGTCGG 6407

QY 366 GCACAGGTGACGCGCCAAACCGGTTGTGAGGGTGCACAGGTCAACGCTGGTGCTA 425
Db 6408 GCACAGGTGACGCGCCAAACCGGTTGTGAGGGTGCACAGGTCAACGCTGGTGCTA 6467

QY 426 CCTCTCCGCTAGGCGCTGTTGGTCCAGCAGAGGCGCGCCCTTGGCGGAGCAAGATG 485
Db 6468 CCTCTCCGCTAGGCGCTGTTGGTCCAGCAGAGGCGCGCCCTTGGCGGAGCAAGATG 6527

QY 486 GCGGTAGGGGGTCTAGTCTGCTCTGTCGCGGGGGTCTGCTGTCACGTTAAGACCCCGG 545
Db 6528 GCGGTAGGGGGTCTAGTCTGCTCTGTCGCGGGGGTCTGCTGTCACGTTAAGACCCCGG 6587

QY 546 GCACAGGCGCGCTCGAAGTAGTCTATCTTCATCTTGCATCTTGCAGTCTAGCGCTCTGCTGCC 605
```

```
Db 6588 GCACAGGCGCGCTCGAAGTAGTCTATCTTGCATCTTGCAGTCTAGCGCTCTGCTGCC 6647
QY 606 ATGCGCGCGCGCAAGCGCGCTCGTATGAGTTGAGTGGGGGACCCCATGGCATGGGGT 665
Db 6648 ATGCGCGCGCGCAAGCGCGCTCGTATGAGTTGAGTGGGGGACCCCATGGCATGGGGT 6707
QY 666 GGGTACGCGCGGAGGCGTACATGCGGCAAAATGCTGAACCTAGAGGGGCTCTCTGAGTA 725
Db 6708 GGGTACGCGCGGAGGCGTACATGCGGCAAAATGCTGAACCTAGAGGGGCTCTCTGAGTA 6767
QY 726 TTCCAAGATATGTAGGCTAGCATCTTCCACCGCGGATGCTGCGCGCACGTAATCGTATA 785
Db 6768 TTCCAAGATATGTAGGCTAGCATCTTCCACCGCGGATGCTGCGCGCACGTAATCGTATA 6827
QY 786 GTTCGTGCGAGGAGCGAGGAGGTGCGGACCCGAGTTGCTACGCGCGGGCTGCTCTGCTC 845
Db 6828 GTTCGTGCGAGGAGCGAGGAGGTGCGGACCCGAGTTGCTACGCGCGGGCTGCTCTGCTC 6887
QY 846 GGAAGACTATCTGCTGAAGATGCGCATGTGAGTTGGATGATATGTTGGACGCTGGAAGA 905
Db 6888 GGAAGACTATCTGCTGAAGATGCGCATGTGAGTTGGATGATATGTTGGACGCTGGAAGA 6947
QY 906 CGTTGAAGCTGCGCTGCTGTGAGACCTACCGGCTCACGCAAGAGGCGGTAGGAGTCCG 965
Db 6948 CGTTGAAGCTGCGCTGCTGTGAGACCTACCGGCTCACGCAAGAGGCGGTAGGAGTCCG 7007
QY 966 GCAGCTTGTGACAGCTCGCGCGGTGACCTGCAAGTCTAGGGCGGAGTAGTCCAGGGTTT 1025
Db 7008 GCAGCTTGTGACAGCTCGCGCGGTGACCTGCAAGTCTAGGGCGGAGTAGTCCAGGGTTT 7067
QY 1026 CCTTGATGATGTCATCTATCTCTGCTCCCTTTTTCACAGCTCGCGTTGAGGACAA 1085
Db 7068 CCTTGATGATGTCATCTATCTCTGCTCCCTTTTTCACAGCTCGCGTTGAGGACAA 7127
QY 1086 ACTCTTCGCGGTCTTTCCAGTACTCTTGGATCGGAAACCCGCTCGGCCCTCCGAAACG 1140
Db 7128 ACTCTTCGCGGTCTTTCCAGTACTCTTGGATCGGAAACCCGCTCGGCCCTCCGAAACG 7182
```

Search completed: July 14, 2005, 14:03:29

Job time : 9291.17 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:35:42 ; Search time 1748.26 Seconds
(without alignments)
4198.742 Million cell updates/sec

Title: US-09-482-682-32
Perfect score: 1240
Sequence: 1 ggatccactctctccgcac.....cagtcacagtcgcaagatct 1240

Scoring table: IDENTITY_NUC
Gapop 10_0 , Gapext 1.0
Searched: 4390206 seqs, 2959870667 residues
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1240	100.0	1240	3	AAA59060 Nucleotid
2	1240	100.0	1240	6	ABA94272 Adenoviru
3	1240	100.0	1240	10	ADB75118 Adenoviru
4	1240	100.0	1240	10	ADF48742 Ad5 tripa
5	1239	99.9	7231	3	AAA59090 Nucleotid
6	1239	99.9	7231	6	ABA94286 Nucleotid
7	1239	99.9	7231	10	ADB75132 Plasmid p
8	1239	99.9	7231	10	ADF48774 Adenoviru
9	1239	99.9	7960	3	AAA59072 Nucleotid
10	1239	99.9	7960	6	ABA94274 Nucleotid
11	1239	99.9	7960	10	ADB75120 Plasmid p
12	1239	99.9	7960	10	ADF48754 Fibre exp
13	1239	99.9	7989	3	AAA59075 Nucleotid
14	1239	99.9	7989	6	ABA94277 Nucleotid
15	1239	99.9	7989	10	ADB75123 Plasmid p
16	1239	99.9	7989	10	ADF48757 Fibre exp
17	1239	99.9	8383	3	AAA59071 Nucleotid
18	1239	99.9	8383	6	ABA94273 Nucleotid
19	1239	99.9	8383	10	ADB75119 Plasmid p
20	1239	99.9	8383	10	ADF48753 Fibre exp

21	1239	99.9	8484	3	AAA59091 Nucleotid
22	1239	99.9	8484	10	ADF48775 Fibre exp
23	1135	91.5	10332	2	AAV33921 Nucleotid
24	1135	91.5	31183	4	AA003963 Adenoviru
25	1135	91.5	31446	3	AA009088 AdPB-bega
26	1135	91.5	31880	12	ADO09305 Ad009305 WTI-F-ade
27	1135	91.5	31976	13	ADR41670 Oncolytic
28	1135	91.5	31976	13	ADR41669 Oncolytic
29	1135	91.5	32026	2	AAT60559 Recombina
30	1135	91.5	32165	3	AAA09092 AdMTV-be
31	1135	91.5	32165	3	AAA14723 Nucleotid
32	1135	91.5	32166	3	AAA09090 AdPSA-bet
33	1135	91.5	32166	4	AA089170 AdRSVPHYD
34	1135	91.5	32167	3	AAA14803 Nucleotid
35	1135	91.5	32167	3	AAZ93332 Partial s
36	1135	91.5	32409	12	ADO36637 Adenovira
37	1135	91.5	32480	3	AAA59055 Nucleotid
38	1135	91.5	32480	6	ABA94267 Adenoviru
39	1135	91.5	32480	10	ADB75113 Adenovira
40	1135	91.5	32480	10	ADF48737 Adenoviru
41	1135	91.5	32681	12	ADO36636 Adenovira
42	1135	91.5	32798	6	ABA97684 Replicati
43	1135	91.5	32802	13	ADR41671 Adr41671 S98-100 (
44	1135	91.5	32886	3	AAA09086 AdRSV-bet
45	1135	91.5	33014	13	ADP79484 Adenoviru

ALIGNMENTS

RESULT 1
AAA59060
ID AAA59060 standard; DNA; 1240 BP.
XX
AC AAA59060;
XX

DT 15-SEP-2003 (revised)
DT 07-NOV-2000 (first entry)
XX

DE Nucleotide sequence of a tripartite leader sequence.

XX Adenovirus; tripartite leader; adenovirus vector particle; gene delivery;
XX ss.
XX Human adenovirus type 5.

XX
XX WO200042208-A1.
XX PD
XX 20-JUL-2000.
XX PF
XX 14-JAN-2000; 2000WO-EP000265.
XX PR
XX 14-JAN-1999; 99US-0115920P.

XX (NOVS) NOVARTIS AG.
XX (NOVS) NOVARTIS-ERFINDUNGEN VERW GES MBH.
XX (SCRI) SCRIPPS RES INST.
XX Nemerow GR, Von Seggern DU, Hallenbeck PL, Stevenson SC;
XX Skripchenko Y;
XX WPI; 2000-476068/41.
XX

XX New nucleic acid comprising an adenovirus tripartite leader nucleotide
XX PT for producing high-capacity and targeted vectors for adenovirus-based
XX gene therapy.
XX Claim 5; Page 180; 212pp; English.

XX The specification describes a nucleic acid molecule comprising an
XX adenovirus (AV) tripartite leader (TPL) nucleotide with a sequence
XX comprising two different TPL exons or three same or different TPL exons.
XX The nucleic acid is used to produce an adenovirus vector particle, CC

CC	deliver an exogenous gene to a target cell, pseudotype recombinant viral
CC	vectors, target an adenovirus vector to a cell, produce a modified
CC	adenovirus, deliver a heterologous gene to an animal and produce a
CC	gutless adenoviral vector particle. The present sequence represents a TPL
CC	sequence, which is used to construct nucleic acid molecules of the
CC	invention.. (Updated on 15-SEP-2003 to standardise OS field)
XX	
SQ	Sequence 1240 BP; 231 A; 327 C; 411 G; 271 T; 0 U; 0 Other;
	Query Match 100.0%; Score 1240; DB 3; Length 1240;
	Best Local Similarity 100.0%; Pred. No. 0;
	Matches 1240; Conservative 0; Mismatches 0; Indels 0; Gaps 0
Qy	1 GGATCCACTCTCTTCCGCATCGCTGTCTGCGAGGCCAGCTGTTGGGTGAGTACTCCCT 60
Db	1 GGATCCACTCTCTTCCGCATCGCTGTCTGCGAGGCCAGCTGTTGGGTGAGTACTCCCT 60
Qy	61 CTGAARAGCGGCATGACTCTTCGGCTAAGATTGTCAGTTTCCAAAACGAGGAGATT 120
Db	61 CTGAARAGCGGCATGACTCTTCGGCTAAGATTGTCAGTTTCCAAAACGAGGAGATT 120
Qy	121 GATATTCACTTCGGCCGCGGTGATGCCCTTCAGGGTGCCCGCATCCATCTGTCAGAAAA 180
Db	121 GATATTCACTTCGGCCGCGGTGATGCCCTTCAGGGTGCCCGCATCCATCTGTCAGAAAA 180
Qy	181 GACAATCTTTTGTTCAGCTTTGGTGGCAAAACGACCCGTAGAGGGCGTTGGACAGCAA 240
Db	181 GACAATCTTTTGTTCAGCTTTGGTGGCAAAACGACCCGTAGAGGGCGTTGGACAGCAA 240
Qy	241 CTTGGCGATGAGCGCAGGGTTGGTTTTTTCGCGATCGGGCGCTCTTGGCCGCGAT 300
Db	241 CTTGGCGATGAGCGCAGGGTTGGTTTTTTCGCGATCGGGCGCTCTTGGCCGCGCAT 300
Qy	301 GTTTAGCTGCATTCGCGCGCAACGCACCGCCATTTCGGGAAAGACGSGTGGTCGCTC 360
Db	301 GTTTAGCTGCATTCGCGCGCAACGCACCGCCATTTCGGGAAAGACGSGTGGTCGCTC 360
Qy	361 GTCGGGCACAGGTGCACGCGCAACCGCGTGTGTGCAGGGTGAACAAGTCAACGCTGGT 420
Db	361 GTCGGGCACAGGTGCACGCGCAACCGCGTGTGTGCAGGGTGAACAAGTCAACGCTGGT 420
Qy	421 GGCTACCTCTCCGCGTAGGCGCTGTTGGTCAGCAGAGGGCGGCCCTTTCGCGGAGCA 480
Db	421 GGCTACCTCTCCGCGTAGGCGCTGTTGGTCAGCAGAGGGCGGCCCTTTCGCGGAGCA 480
Qy	481 GAATGGCGGTAGGGGGTCTAGTGTGGTCTCGTTCGGGGGGTCTGCGTCCACGTAAGAC 540
Db	481 GAATGGCGGTAGGGGGTCTAGTGTGGTCTCGTTCGGGGGGTCTGCGTCCACGTAAGAC 540
Qy	541 CCCGGGACGAGCGCGCGTGTGAGTAGTCTATCTTGATCTCTTGCAGAGTCTAGGCGCTG 600
Db	541 CCCGGGACGAGCGCGCGTGTGAGTAGTCTATCTTGATCTCTTGCAGAGTCTAGGCGCTG 600
Qy	601 CTGCCATCGCGCGGCGCAAGCGCGCTCGTATGGGTTGAGTGGGGACCCCATGGCAT 660
Db	601 CTGCCATCGCGCGGCGCAAGCGCGCTCGTATGGGTTGAGTGGGGACCCCATGGCAT 660
Qy	661 GGGGTGGGTGAGCGCGGAGGCGGTATCATCGCGAAATGTCGTAAACGTAGAGGGGCTCTCT 720
Db	661 GGGGTGGGTGAGCGCGGAGGCGGTATCATCGCGAAATGTCGTAAACGTAGAGGGGCTCTCT 720
Qy	721 GAGTATTCGAAGATATGTAAGGTAGCATCTTCCACGGCGGATGCTGGCGCGCACGTAATC 780
Db	721 GAGTATTCGAAGATATGTAAGGTAGCATCTTCCACGGCGGATGCTGGCGCGCACGTAATC 780
Qy	781 GTATAGTTTCGTGCGAGGAGCGAGGAGTTCGGGACCGAGGTTGCTACGGGGGGCTGCTC 840
Db	781 GTATAGTTTCGTGCGAGGAGCGAGGAGTTCGGGACCGAGGTTGCTACGGGGGGCTGCTC 840
Qy	841 TGCTCGGAAGACTATCTCCCTGAAGATGGCATGTGAGTTGGATGATA TGGTTGGAACGCTG 900
Db	841 TGCTCGGAAGACTATCTCCCTGAAGATGGCATGTGAGTTGGATGATA TGGTTGGAACGCTG 900

CC degenerative disease, retinitis pigmentosa, Stargardt's disease, diabetic
CC retinopathies, retinal vascularizations, and retinoblastoma, of a mammal
CC preferably human. The AAV comprises a fiber protein that specifically or
CC selectively binds to receptors that are expressed on cells (preferably
CC photoreceptors in the eye). Preferably, the recombinant virus comprise a
CC fiber protein from an adenovirus type D subgroup or is a chimeric protein
CC containing a portion of the N-terminus of an adenovirus type 2 or type 5
CC penton, and the therapeutic product is a tropic factor, an anti-
CC apoptotic factor, gene encoding a rhodopsin protein, a wild-type
CC stargardt disease gene (STGD1), an anti-cancer agent and a protein that
CC regulates expression of a photoreceptor specific gene product. The viral
CC nucleic acid of AAV comprises ITRS and packaging signal derived from AAV
CC subgroup B or C, especially an AV type 2 or type 5. AAV is also useful
CC for targeted gene therapy, where the vector comprises an AV type 37 fiber
CC protein or its portion, and selectively transduces photoreceptors and
CC delivers a gene product encoded by AAV. The present sequence represents a
CC adenovirus 5 tripartite leader (TPL) nucleotide sequence. (Updated on 07-
XX AUG-2003 to correct OS field.)

Sequence 1240 BP; 231 A; 327 C; 411 G; 271 T; 0 U; 0 Other;

Query Match 100.0%; Score 1240; DB 6; Length 1240;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATCCACTCTCTCCGATCCTCTCTCGAGGCGCAGCTGTGGGGTGAATCTCCCT 60
DB 1 GGATCCACTCTCTCCGATCCTCTCTCGAGGCGCAGCTGTGGGGTGAATCTCCCT 60
QY 61 CTGAAAGCGGCATGACTTTCGCGTAAAGTGTCTTCCAAAACGAGGAGTTT 120
DB 61 CTGAAAGCGGCATGACTTTCGCGTAAAGTGTCTTCCAAAACGAGGAGTTT 120
QY 121 GATATTCACCTGGCGCGGCTGATCCCTTTGAGGGTGGCGCATCATCTGTGTGAGAAA 180
DB 121 GATATTCACCTGGCGCGGCTGATCCCTTTGAGGGTGGCGCATCATCTGTGTGAGAAA 180
QY 181 GACAACTCTTTTGTGTCAAGCTTGTGGTGGCAACGACCGGTAGAGGGGTGGACAGCAA 240
DB 181 GACAACTCTTTTGTGTCAAGCTTGTGGTGGCAACGACCGGTAGAGGGGTGGACAGCAA 240
QY 241 CTGCGGATGAGCGAGGGTTGGTTTGTGCGGATCGCGCGCTCTTGGCGCGCAT 300
DB 241 CTGCGGATGAGCGAGGGTTGGTTTGTGCGGATCGCGCGCTCTTGGCGCGCAT 300
QY 301 GTTTAGCTGCAGTATTCGCGCGCAACGACCGCATTCGGGAAGACGGTGTGCGCTC 360
DB 301 GTTTAGCTGCAGTATTCGCGCGCAACGACCGCATTCGGGAAGACGGTGTGCGCTC 360
QY 361 GTGCGGCACAGGTGCACGCGCAACCGCGGTGTGCGAGGGTGCAGAGGTCAACGCTGGT 420
DB 361 GTGCGGCACAGGTGCACGCGCAACCGCGGTGTGCGAGGGTGCAGAGGTCAACGCTGGT 420
QY 421 GGTACTCTCTCCGATGAGCGTCTGTGTGTCAGAGAGGCGCGCTCTTGGCGGAGCA 480
DB 421 GGTACTCTCTCCGATGAGCGTCTGTGTGTCAGAGAGGCGCGCTCTTGGCGGAGCA 480
QY 481 GAATGCGGTAGGGGTCTAGTGTGCTCTGTCGCGGGGGTCTGCGTCCACGGTAAAGAC 540
DB 481 GAATGCGGTAGGGGTCTAGTGTGCTCTGTCGCGGGGGTCTGCGTCCACGGTAAAGAC 540
QY 541 CCCGGGACAGCGCGCTGCAAGTAGTCTATCTTGCATCTTGCAGAGTCTAGCGCTG 600
DB 541 CCCGGGACAGCGCGCTGCAAGTAGTCTATCTTGCATCTTGCAGAGTCTAGCGCTG 600
QY 601 CTGCCATGCGCGGCGCAAGCGCGCTCGTATGAGTGGGGGACCCCATGGCAT 660
DB 601 CTGCCATGCGCGGCGCAAGCGCGCTCGTATGAGTGGGGGACCCCATGGCAT 660
QY 661 GGGGTGGTGAAGCGGAGCGGTATGCGCGCAATGTCTGTAACGTAAGAGGGGTCTCT 720
DB 661 GGGGTGGTGAAGCGGAGCGGTATGCGCGCAATGTCTGTAACGTAAGAGGGGTCTCT 720

QY 721 GAGTATTTCCAAAGATATGTAGGGTAGCATCTTCCACCGCGGATGCTGGCGCGCACTAATC 780
DB 721 GAGTATTTCCAAAGATATGTAGGGTAGCATCTTCCACCGCGGATGCTGGCGCGCACTAATC 780
QY 781 GTATAGTTCTGCGAGGGAGCGAGAGGTCCGGACCGAGGTTGCTACCGGGGGCTGCTC 840
DB 781 GTATAGTTCTGCGAGGGAGCGAGAGGTCCGGACCGAGGTTGCTACCGGGGGCTGCTC 840
QY 841 TGCTCGGAAGACTATCTGCTCTGAAAGATGGCATGTGAGTTGGATGATATGGTTGGACGCTG 900
DB 841 TGCTCGGAAGACTATCTGCTCTGAAAGATGGCATGTGAGTTGGATGATATGGTTGGACGCTG 900
QY 901 GAAGACGTTGAAGCTGGGCTCTGTGAGACCTACCGCTGACGACGAGGAGGCGTAGGA 960
DB 901 GAAGACGTTGAAGCTGGGCTCTGTGAGACCTACCGCTGACGACGAGGAGGCGTAGGA 960
QY 961 GTGCGGACGCTGTTGTTGACGAGCTGGCGGTGACCTGTAGGCGCGCAGTAGTCTCCAG 1020
DB 961 GTGCGGACGCTGTTGTTGACGAGCTGGCGGTGACCTGTAGGCGCGCAGTAGTCTCCAG 1020
QY 1021 GGTTCCTTGTATGATGTATCTTATCTCTGCTCCCTTTTTCACAGCTCGCGGTTGAG 1080
DB 1021 GGTTCCTTGTATGATGTATCTTATCTCTGCTCCCTTTTTCACAGCTCGCGGTTGAG 1080
QY 1081 GACAACTCTTTCGCGGCTCTTTCAGATCTCTTGGATCGGAAACCGCTCGGCTCCGAAACG 1140
DB 1081 GACAACTCTTTCGCGGCTCTTTCAGATCTCTTGGATCGGAAACCGCTCGGCTCCGAAACG 1140
QY 1141 AGATCCGTTACTCCGCGCGGAGGACCTGAGCGAGTCCGCGATCGACCGGATCGGAAACG 1200
DB 1141 AGATCCGTTACTCCGCGCGGAGGACCTGAGCGAGTCCGCGATCGACCGGATCGGAAACG 1200
QY 1201 TCTCGAAGAAAGCGCTTAAACCAAGTCACAGTCGCAAGATCT 1240
DB 1201 TCTCGAAGAAAGCGCTTAAACCAAGTCACAGTCGCAAGATCT 1240

RESULT 3

ADB75118

ID ADB75118 standard; DNA; 1240 BP.

XX ADB75118;

XX 04-DEC-2003 (first entry)

XX Adenovirus type 5 tripartite leader sequence #2.

XX ophthalmological; antiinflammatory; antidiabetic; gene therapy;

XX adenovirus inverted terminal repeat sequence;

XX adenovirus packaging signal; photoreceptor-specific promoter;

XX adenovirus type 37; adenovirus type D serotype; adenovirus type 2;

XX adenovirus type 5; photoreceptor; trophic factor; anti-apoptotic factor;

XX rhodopsin; wild-type Stargardt disease gene; STGD1; anti-cancer agent;

XX retinal degenerative disease; retinitis pigmentosa; Stargardt's disease;

XX gyrate atrophy; macular dystrophy; retinoblastoma;

XX photoreceptor-restricted transgene expression;

XX recombinant adenovirus vector; adenovirus type 5; Ad5;

XX tripartite leader sequence; TPL; ds.

XX Human adenovirus type 5.

XX US2002193327-A1.

XX 19-DEC-2002.

XX 01-MAY-2001; 2001US-00847101.

XX 01-MAY-2000; 2000US-00562934.

XX (SCRI) SCRIPPS RES INST.

XX Nemerow GR, Von Seggern DJ, Friedlander M;

XX WPI; 2003-657234/62.
XX
PT Novel nucleic acids comprising adenovirus inverted terminal repeat
PT sequences, adenovirus packaging signals operatively linked to the
PT sequences and photoreceptor-specific promoters, useful for treating
PT retinitis pigmentosa.
XX
PS Example 3; Page 78; 106pp; English.
XX
CC The invention describes an isolated nucleic acid (I) comprising
CC adenovirus inverted terminal repeat sequence, an adenovirus packaging
CC signal operatively linked to the sequence, and a photoreceptor-specific
CC promoter. A Recombinant adenovirus vector (II) comprising (I) is useful
CC for targeted delivery of a gene product to the eye of a mammal which
CC involves administering (II) that comprises heterologous DNA encoding the
CC gene product or resulting in expression of the gene product, where the
CC recombinant virus comprises a fibre protein that specifically or
CC selectively binds to receptors that are expressed on cells which are
CC photoreceptors, in the eye. The recombinant virus comprises a fibre
CC protein which is an adenovirus type 37, from an adenovirus type D
CC serotype. The fibre is a chimeric protein containing a sufficient portion
CC of the N-terminus of an adenovirus type 2 or type 5 fibre protein for
CC interaction with an adenovirus type 2 or type 5 penton, and a sufficient
CC portion of an adenovirus serotype D knob portion of the fiber for
CC selective binding to photoreceptors in the eye of a mammal. The
CC encapsulated nucleic acid comprises a photoreceptor-specific promoter
CC operatively linked to a nucleic acid comprising the therapeutic product
CC which is chosen from trophic factor, anti-apoptotic factor, gene encoding
CC a rhodopsin protein, wild-type Stargardt disease gene (STDG1), an anti-
CC cancer agent and a protein that regulates expression of a photoreceptor-
CC specific gene product. The delivery is effected for treatment of an
CC ocular disease such as retinal degenerative disease e.g., retinitis
CC pigmentosa, Stargardt's disease, diabetic retinopathies, retinal
CC vascularisation, choroideremia, gyrate atrophy or macular dystrophy or
CC retinoblastoma inherited and acquired retinal and neovascular
CC degenerative diseases. The viral nucleic acid comprises an adenovirus
CC inverted terminal repeat (ITR) sequences, and an adenovirus packaging
CC signal operatively linked to the sequence. The ITRs and packaging signal
CC are derived from an adenovirus serotype B or C, or adenovirus type 2 or
CC 5. The viral nucleic acid further comprises a photoreceptor-specific
CC promoter. (II) includes photoreceptor promoters providing a means not
CC only for specific targeting of expression in these cells, but also for
CC photoreceptor-restricted transgene expression. This sequence represents a
CC TPL (tripartite leader sequence) from the adenovirus type 5 genome, used
CC to enhance the expression of complementing adenoviral proteins.
XX
SQ Sequence 1240 BP; 231 A; 327 C; 411 G; 271 T; 0 U; 0 Other;

Query Match 100.0%; Score 1240; DB 10; Length 1240;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATCCACTCTCTCCGATGCTCTCTGCGAGGCGCAGCTGTGGGGTGAGTACTCCCT 60
DB 1 GGATCCACTCTCTCCGATGCTCTCTGCGAGGCGCAGCTGTGGGGTGAGTACTCCCT 60
QY 61 CTGAAAGCGGCGCATGACTTCTCGCGCTAAGATTGTTCAGTTTCCAAAAACGAGGAGATT 120
DB 61 CTGAAAGCGGCGCATGACTTCTCGCGCTAAGATTGTTCAGTTTCCAAAAACGAGGAGATT 120
QY 121 GATATTACATGGCCCGCGGTGATGCTCTTGGGGTGGCGGATCCATCTGGTCAAAAA 180
DB 121 GATATTACATGGCCCGCGGTGATGCTCTTGGGGTGGCGGATCCATCTGGTCAAAAA 180
QY 181 GACAATCTTTTGTGTCAAGCTTCGCGCAACCAACCGTAGAGGGGTGGACAGCAA 240
DB 181 GACAATCTTTTGTGTCAAGCTTCGCGCAACCAACCGTAGAGGGGTGGACAGCAA 240
QY 241 CTTGGCGATGAGCGCAGGGTTGGTTTTGTGCGGATCGGCGGCTCTTTGGCGCGAT 300
DB 241 CTTGGCGATGAGCGCAGGGTTGGTTTTGTGCGGATCGGCGGCTCTTTGGCGCGAT 300

QY 301 GTTTAGCTGCACGTTATTCGCGCGCAACGACCGCCATTTCGGGAAAGAGGTGTGGCGCTC 360
DB 301 GTTTAGCTGCACGTTATTCGCGCGCAACGACCGCCATTTCGGGAAAGAGGTGTGGCGCTC 360
QY 361 GTCGGGCAACAGGTGACGCGGCCAAACCGCGGTTGTGAGGGGTGACAAGGTCAACGCTGGT 420
DB 361 GTCGGGCAACAGGTGACGCGGCCAAACCGCGGTTGTGAGGGGTGACAAGGTCAACGCTGGT 420
QY 421 GGCTACCTCTCCGCGTAGGCGCTGCTGGTGTCCAGCAGAGGCGCGCGCTTGGCGGAGCA 480
DB 421 GGCTACCTCTCCGCGTAGGCGCTGCTGGTGTCCAGCAGAGGCGCGCGCTTGGCGGAGCA 480
QY 481 GAATGGCGGTAGGGGTTAGCTGCGCTCTGTCGGGGGGTCTGCGTCCACGGTAAGAC 540
DB 481 GAATGGCGGTAGGGGTTAGCTGCGCTCTGTCGGGGGGTCTGCGTCCACGGTAAGAC 540
QY 541 CCCGGGCAAGAGGCGCGCTGCAAGTAGTCTATCTTGCATCTTTCGAAGTCTAGCGCGCTG 600
DB 541 CCCGGGCAAGAGGCGCGCTGCAAGTAGTCTATCTTGCATCTTTCGAAGTCTAGCGCGCTG 600
QY 601 CTGCCATGCGCGGCGCAAGCGCGCTCGTATGCGTGTGAGTGGGGGACCCCATGGCAT 660
DB 601 CTGCCATGCGCGGCGCAAGCGCGCTCGTATGCGTGTGAGTGGGGGACCCCATGGCAT 660
QY 661 GGGTGGGTGAGCGCGGAGCGGTACATGCCCAATCTCGTAAACGTAGAGGGGCTCTCT 720
DB 661 GGGTGGGTGAGCGCGGAGCGGTACATGCCCAATCTCGTAAACGTAGAGGGGCTCTCT 720
QY 721 GAGTATTCGAAGATATGAGGTAGCATCTTCCACCGGATGCTGGCGCGCATATC 780
DB 721 GAGTATTCGAAGATATGAGGTAGCATCTTCCACCGGATGCTGGCGCGCATATC 780
QY 781 GTATAGTTCGTGCGAGGAGCGAGGTGCGGACCGAGGTGTGACGGGCGGCTGCTC 840
DB 781 GTATAGTTCGTGCGAGGAGCGAGGTGCGGACCGAGGTGTGACGGGCGGCTGCTC 840
QY 841 TGCTCGAAGACTATCTGCTGAGATGGATGCTGAGTGGATGATATGTTGGACGCTG 900
DB 841 TGCTCGAAGACTATCTGCTGAGATGGATGCTGAGTGGATGATATGTTGGACGCTG 900
QY 901 GAAGACGTTGAAGCTGCGCTCTGTGAGACCTTACCGCGTCACGCAAGAGGCGGTAGGA 960
DB 901 GAAGACGTTGAAGCTGCGCTCTGTGAGACCTTACCGCGTCACGCAAGAGGCGGTAGGA 960
QY 961 GTCGCGCAGCTGTTGTAACAGCTGCGGCGTACCTGCACTGTTAGGGCGAGTAGTCCAG 1020
DB 961 GTCGCGCAGCTGTTGTAACAGCTGCGGCGTACCTGCACTGTTAGGGCGAGTAGTCCAG 1020
QY 1021 GGTTCCTTGTGATGATGTCATCTTATCTGTCCTCTTTTTCACAGCTCGCGGTTGAG 1080
DB 1021 GGTTCCTTGTGATGATGTCATCTTATCTGTCCTCTTTTTCACAGCTCGCGGTTGAG 1080
QY 1081 GACAAACTCTTTCGCGGCTTTTCCAGTACTCTTGGATCGGAAACCCGTCGGCTCCGAACG 1140
DB 1081 GACAAACTCTTTCGCGGCTTTTCCAGTACTCTTGGATCGGAAACCCGTCGGCTCCGAACG 1140
QY 1141 AGATCCGTACTTCCCGCGCGAGGAGCTGAGCGAGTCCGCACTGACCGGATCGGAAACCC 1200
DB 1141 AGATCCGTACTTCCCGCGCGAGGAGCTGAGCGAGTCCGCACTGACCGGATCGGAAACCC 1200
QY 1201 TCTCGAAGAGGCGTCTAAACAGTCCAGTCCAGATCT 1240
DB 1201 TCTCGAAGAGGCGTCTAAACAGTCCAGTCCAGATCT 1240

RESULT 4
ADF48742
ID ADF48742 standard; DNA; 1240 BP.
XX
XX ADF48742;
XX AC
XX DT 12-FEB-2004 (first entry)
XX


```
QY 1141 AGATCCGTACTCCGCGCGGAGGACCTGAGGAGTCCGATCGACCGGATCGGAAAC 1200
Db 1141 AGATCCGTACTCCGCGCGGAGGAGCTGAGCGAGTCCGATCGACCGGATCGGAAAC 1200
QY 1201 TCTCAGAAAGCGCTTAACAGTCAACAGTCAACAGTCAACAGTCAACAGTCT 1240
Db 1201 TCTCAGAAAGCGCTTAACAGTCAACAGTCAACAGTCAACAGTCAACAGTCT 1240

RESULT 5
AAAS9090
ID AAAS9090 standard; DNA; 7231 BP.
XX
AC AAAS9090;
XX
DT 07-NOV-2000 (first entry)
XX
DE Nucleotide sequence of plasmid pDV80.
XX
KW Adenovirus; tripartite leader; adenovirus vector particle; gene delivery;
KW ss.
XX
OS Synthetic.
XX
PN WO200042208-A1.
XX
PD 20-JUL-2000.
XX
PF 14-JAN-2000; 2000WO-EP000265.
XX
PR 14-JAN-1999; 99US-0115920P.
XX
PA (NOVS ) NOVARTIS AG.
PA (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.
PA (SCRI ) SCRIPPS RES INST.
XX
XX
PI Nemerow GR, Von Seggern DJ, Hallenbeck PL, Stevenson SC;
PI Skripchenko Y;
XX
XX WPI; 2000-476068/41.
XX
PT New nucleic acid comprising an adenovirus tripartite leader nucleotide
PT for producing high-capacity and targeted vectors for adenovirus-based
PT gene therapy.
XX
PS Claim 10; Page 200-201; 212pp; English.
XX
CC The specification describes a nucleic acid molecule comprising an
CC adenovirus (AV) tripartite leader (TPL) nucleotide with a sequence
CC comprising two different TPL exons or three same or different TPL exons.
CC The nucleic acid is used to produce an adenovirus vector particle,
CC deliver an exogenous gene to a target cell, pseudotype recombinant viral
CC vectors, target an adenovirus vector to a cell, produce a modified
CC adenovirus, deliver a heterologous gene to an animal and produce a
CC gutless adenoviral vector particle. The present sequence represents
CC plasmid pDV80, which contains a TPL
XX
SQ Sequence 7231 BP; 1751 A; 1807 C; 1869 G; 1804 T; 0 U; 0 Other;

Query Match 99.9%; Score 1239; DB 3; Length 7231;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATCCACTCTCTCCGATCGCTGTCGAGGCGCAGCTCTTGGGTGAGTACTCCCT 60
Db 849 GGATCCACTCTCTCCGATCGCTGTCGAGGCGCAGCTGTTGGGTGAGTACTCCCT 908
QY 61 CTGAAAGCGGCATGACTTTCGCTAAGATTGTCAAGTTTCCAAAAACGAGGAGATT 120
Db 909 CTGAAAGCGGCATGACTTTCGCTAAGATTGTCAAGTTTCCAAAAACGAGGAGATT 968
QY 121 GATATTACCTGGCCGCGGTGATGCTTTGAGGGTGGCGGCATCCATCTGTCAGAAA 180
```

```
Db 969 GATATTACCTGGCCGCGGTGATGCTTTGAGGGTGGCGGCATCCATCTGTCAGAAA 1028
QY 181 GACAATCTTTTGTGTCAAGCTTTGGTGGCAACGACCCGCTAGAGGGCGTTGGACAGCA 240
Db 1029 GACAATCTTTTGTGTCAAGCTTTGGTGGCAACGACCCGCTAGAGGGCGTTGGACAGCA 1088
QY 241 CTTGGCGATGAGCGCAGGGTTTGGTTTGTTCGGATCGCGCGCTCTTGGCGCGGAT 300
Db 1089 CTTGGCGATGAGCGCAGGGTTTGGTTTGTTCGGATCGCGCGCTCTTGGCGCGGAT 1148
QY 301 GTTTAGCTGCACGTATTTCGGCGCAACGACCGCCATTTCGGGAAAGACGGTGTGCGCTC 360
Db 1149 GTTTAGCTGCACGTATTTCGGCGCAACGACCGCCATTTCGGGAAAGACGGTGTGCGCTC 1208
QY 361 GTCGGGCAACAGGTGCAACCGCGCTGTGAGGGTGAACAAGTCAACGCTGGT 420
Db 1209 GTCGGGCAACAGGTGCAACCGCGCTGTGAGGGTGAACAAGTCAACGCTGGT 1268
QY 421 GGCTACCTCTCCGGTAGGGCTCTAGTCTCGTCCGGGGGTCTCGTCCACGGTAAAGAC 480
Db 1269 GGCTACCTCTCCGGTAGGGCTCTAGTCTCGTCCGGGGGTCTCGTCCACGGTAAAGAC 1328
QY 481 GAATGGCGTAGGGGTCTAGTCTCGTCCGGGGGTCTCGTCCACGGTAAAGAC 540
Db 1329 GAATGGCGTAGGGGTCTAGTCTCGTCCGGGGGTCTCGTCCACGGTAAAGAC 1388
QY 541 CCGGGCAGCAGCGCGCTGAAAGTAGTCTATCTTGCATCTCTTGCAGTCTAGCGCTG 600
Db 1389 CCGGGCAGCAGCGCGCTGAAAGTAGTCTATCTTGCATCTCTTGCAGTCTAGCGCTG 1448
QY 601 CTGCCATCGCGGCGGCAAGCGCGCTCGTATGGGTGAGTGGGGACCCCATGGCAT 660
Db 1449 CTGCCATCGCGGCGGCAAGCGCGCTCGTATGGGTGAGTGGGGACCCCATGGCAT 1508
QY 661 GGGGTGGGTGAGCGGAGCGGTACATGCGCAAAATGTCGTAACGTAGAGGGGCTCTCT 720
Db 1509 GGGGTGGGTGAGCGGAGCGGTACATGCGCAAAATGTCGTAACGTAGAGGGGCTCTCT 1568
QY 721 GAGTATTCGAAGATATGAGGGTAGCTCTTCCACCGCGGATGTGGCGCGCACGTAATC 780
Db 1569 GAGTATTCGAAGATATGAGGGTAGCTCTTCCACCGCGGATGTGGCGCGCACGTAATC 1628
QY 781 GTATAGTTCTGCGAGGAGCGAGAGGTCCGGAACGAGTTGCTACGGGCGGCTGCTC 840
Db 1629 GTATAGTTCTGCGAGGAGCGAGAGGTCCGGAACGAGTTGCTACGGGCGGCTGCTC 1688
QY 841 TGCTCGGAAGACTATCTGCCTGAAGATGGCATGTGAGTTGGATGATATGTTGGACGCTG 900
Db 1689 TGCTCGGAAGACTATCTGCCTGAAGATGGCATGTGAGTTGGATGATATGTTGGACGCTG 1748
QY 901 GAAGACGTTGAAGCTGGCGTCTGTGAGACCTTACCGCGTACGACGACGAAGAGGGTAGGA 960
Db 1749 GAAGACGTTGAAGCTGGCGTCTGTGAGACCTTACCGCGTACGACGACGAAGAGGGTAGGA 1808
QY 961 GTCGCGCAGCTTGTGACCAGCTCGCGGTGACCTGCACGCTTAGGGCGCAGTAGTCCAG 1020
Db 1809 GTCGCGCAGCTTGTGACCAGCTCGCGGTGACCTGCACGCTTAGGGCGCAGTAGTCCAG 1868
QY 1021 GGTTCCTTGATGATGTCATCTTATCTGTCCTCTTTTTCACAGCTCGCGGTTGAG 1080
Db 1869 GGTTCCTTGATGATGTCATCTTATCTGTCCTCTTTTTCACAGCTCGCGGTTGAG 1928
QY 1081 GACAAACTCTTCGCGGTCTTTTCCAGTACTCTTGGATCGGAAACCGCTCGGCTCCGAAAC 1140
Db 1929 GACAAACTCTTCGCGGTCTTTTCCAGTACTCTTGGATCGGAAACCGCTCGGCTCCGAAAC 1988
QY 1141 AGATCCGTACTCCGCGCGGAGGACCTGAGCGAGTCCGATCGACCGGATCGGAAAC 1200
Db 1989 AGATCCGTACTCCGCGCGGAGGACCTGAGCGAGTCCGATCGACCGGATCGGAAAC 2048
QY 1201 TCTCAGAAAGCGCTCTAAACAGTCAACAGTCAACAGTCAACAGTCTGTCAGAAA 1239
```


Db 2049 TCTCGAGAAAGCGTCTTAACCAAGTCACAGTCGCAAGATC 2087

RESULT 6

ABA94286
ID ABA94286 standard; DNA; 7231 BP.

XX ABA94286;
AC ABA94286;

DT 13-MAR-2002 (first entry)

XX Nucleotide sequence of expression plasmid pDV80.

XX Adenovirus; inverter terminal repeat sequence; ITRS; ocular disease;
KW fiber protein; photoreceptor; rhodopsin; stargardt disease gene; STDG1;
KW ophthalmological; antiinflammatory; antidiabetic; cytostatic;
KW gene therapy; fiber protein; ss.

XX Synthetic.

XX WO200183729-A2.

XX 08-NOV-2001.

XX 30-APR-2001; 2001WO-EP004863.

XX 01-MAY-2000; 2000US-00562934.

XX (NOVS) NOVARTIS AG.

XX (SCRI) SCRIPPS RES INST.

XX (NEME/) NEMEROW G R.

XX (VSEG/) VON SEGGERN D J.

XX (FRIE/) FRIEDLANDER M.

PI Nemerow GR, Von Seggern DJ, Friedlander M;

XX WPI; 2002-082846/11.

XX Polynucleotide for making vectors, useful for treating ocular diseases,
PT e.g., retinitis pigmentosa, comprises adenovirus inverter terminal repeat
PT sequences, packaging signal and photoreceptor-specific promoter.

XX Example 8; Page 146-148; 149pp; English.

XX The invention provides an isolated polynucleotide comprising adenovirus
CC (AV) inverter terminal repeat sequences (ITRS), AV packaging signal
CC operatively linked to ITRS and a photoreceptor-specific promoter. A
CC recombinant AV vector (AVV) comprising the polynucleotide is useful for
CC targeted delivery of a gene product to the eye (especially to the
CC vitreous cavity), for treating an ocular disease, e.g., retinal
CC degenerative disease, retinitis pigmentosa, Stargardt's disease, diabetic
CC retinopathies, retinal vascularizations, and retinoblastoma, of a mammal
CC preferably human. The AVV comprises a fiber protein that specifically or
CC selectively binds to receptors that are expressed on cells (preferably
CC photoreceptors in the eye). Preferably, the recombinant virus comprise a
CC fiber protein from an adenovirus type D subgroup or is a chimeric protein
CC containing a portion of the N-terminus of an adenovirus type 2 or type 5
CC penton, and the therapeutic product is a trophic factor, an anti-
CC apoptotic factor, gene encoding a rhodopsin protein, a wild-type
CC stargardt disease gene (STDG1), an anti-cancer agent and a protein that
CC regulates expression of a photoreceptor specific gene product. The viral
CC nucleic acid of AAV comprises ITRS and packaging signal derived from AAV
CC subgroup B or C, especially an AV type 2 or type 5. AAV is also useful
CC for targeted gene therapy, where the vector comprises an AV type 37 fiber
CC protein or its portion, and selectively transduces photoreceptors and
CC delivers a gene product encoded by AAV. The present sequence represents
CC the nucleotide sequence of plasmid pDV80, an expression plasmid for
CC adenoviral 37 fiber protein

XX Sequence 7231 BP; 1751 A; 1807 C; 1869 G; 1804 T; 0 U; 0 Other;

Query Match 99.9%;

Best Local Similarity 100.0%; Pred. No. 0;

Score 1239; DB 6; Length 7231;

	Matches 1239;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY 1	GGATCCACTCTCTTCCCGCATCGTGTCTGCGAGGCGCAGCTGTGGGGTGTAGTACTCCCT				60
DB	GGATCCACTCTCTTCCCGCATCGTGTCTGCGAGGCGCAGCTGTGGGGTGTAGTACTCCCT				908
QY 61	CTGAAAAGCGGGCATGACTTCTGCGCTAAGATTGTCAAGTTTCCAAAAACGAGGAGGATTT				120
DB	CTGAAAAGCGGGCATGACTTCTGCGCTAAGATTGTCAAGTTTCCAAAAACGAGGAGGATTT				968
QY 121	GATATTACCTGGCCCGCGGTGATGCTTTCAGGGTGGCCGATCCATCTGTCAGAAAA				180
DB	GATATTACCTGGCCCGCGGTGATGCTTTCAGGGTGGCCGATCCATCTGTCAGAAAA				1028
QY 181	GACAACTCTTTTGTGTCAAGCTTGGTGGCAAAACGACCCGTAGAGGGCGTTGGACAGCAA				240
DB	GACAACTCTTTTGTGTCAAGCTTGGTGGCAAAACGACCCGTAGAGGGCGTTGGACAGCAA				1088
QY 241	CTTGGCGATGAGCGCAGGGTTGGTTTGTCTCGGATCGCGCGCTCTTGGCCGCGAT				300
DB	CTTGGCGATGAGCGCAGGGTTGGTTTGTCTCGGATCGCGCGCTCTTGGCCGCGAT				1148
QY 301	GTTTAGCTGCACGTATTTCGCGCGCAACGACCGCCATTTCGGGAAGACGCTGTCGCTC				360
DB	GTTTAGCTGCACGTATTTCGCGCGCAACGACCGCCATTTCGGGAAGACGCTGTCGCTC				1208
QY 361	GTCGGGCAACAGGTGACGCGCCAAACGCGGTTGTGAGGGTGCACAGGTCAACGCTGGT				420
DB	GTCGGGCAACAGGTGACGCGCCAAACGCGGTTGTGAGGGTGCACAGGTCAACGCTGGT				1268
QY 421	GGTAACTCTCCCGTAGGCGCTCGTTGGTCCAGCAGAGCGCGCGCTTTCGCGCAGCA				480
DB	GGTAACTCTCCCGTAGGCGCTCGTTGGTCCAGCAGAGCGCGCGCTTTCGCGCAGCA				1328
QY 481	GAATGGCGGTAGGGGTTCTAGCTCGCTCTCGTCGGGGGCTCGCTCCAGCGTAAGAC				540
DB	GAATGGCGGTAGGGGTTCTAGCTCGCTCTCGTCGGGGGCTCGCTCCAGCGTAAGAC				1388
QY 541	CCCCGGCAGCAGGCGCGCTCGAAGTAGTCTATCTTGCATCTTGCAGTCTAGCGCCTG				600
DB	CCCCGGCAGCAGGCGCGCTCGAAGTAGTCTATCTTGCATCTTGCAGTCTAGCGCCTG				1448
QY 601	CTGCCATCTCGCGCGGCAAGCGCGCTCGTATGGGTTGAGTGGGGACCCCATGGCAT				660
DB	CTGCCATCTCGCGCGGCAAGCGCGCTCGTATGGGTTGAGTGGGGACCCCATGGCAT				1508
QY 661	GGGTGGGTGAGCGCGGAGGCTACATCCGCCAAATGTCGTAAGTAGAGGGGCTCTCT				720
DB	GGGTGGGTGAGCGCGGAGGCTACATCCGCCAAATGTCGTAAGTAGAGGGGCTCTCT				1568
QY 721	GAGTATTCGAAGATATGTAGGGTAGCATCTTCCACCGCGATGCTGGCGCGCAGTAATC				780
DB	GAGTATTCGAAGATATGTAGGGTAGCATCTTCCACCGCGATGCTGGCGCGCAGTAATC				1628
QY 781	GTATAGTTCGTGCGAGGAGCGAGGAGTCCGGACCCAGGTTGCTACGGGGGGTGTCTC				840
DB	GTATAGTTCGTGCGAGGAGCGAGGAGTCCGGACCCAGGTTGCTACGGGGGGTGTCTC				1688
QY 841	TGCTCGGAAGACTATCTGCTTGAAGATGGCATGTGAGTTGGATGATATGTGTGACCGTG				900
DB	TGCTCGGAAGACTATCTGCTTGAAGATGGCATGTGAGTTGGATGATATGTGTGACCGTG				1748
QY 901	GAACAGCTTGAAGCTGGCTCTGTGAGACCTTACCGCTACGCGACGAGGAGGCTAGGA				960
DB	GAACAGCTTGAAGCTGGCTCTGTGAGACCTTACCGCTACGCGACGAGGAGGCTAGGA				1808
QY 961	GTCGCGCAGCTTGTGTGACCGAGTCCGCGGTGACCTGACGCTTAGGGGCGCAGTAGTCAG				1020
DB	GTCGCGCAGCTTGTGTGACCGAGTCCGCGGTGACCTGACGCTTAGGGGCGCAGTAGTCAG				1868
QY 1021	GGTTTCTTGTGATGATCATCTTATCTGTGCTTTTTCACAGCTCGCGGTTGAG				1080
DB	GGTTTCTTGTGATGATCATCTTATCTGTGCTTTTTCACAGCTCGCGGTTGAG				1928

QY 1081 GACAAACTCTTCGGCGTCTTTCCAGTACTCTTGGATCGGAACACCGTCGGCTCGGAACG 1140
 Db 1929 GACAAACTCTTCGGCGTCTTTCCAGTACTCTTGGATCGGAACACCGTCGGCTCGGAACG 1988
 QY 1141 AGATCCGCTACTCCGCGCGGAGGACCTGAGCGAGTCCGCATCGACCGGATCGGAACAC 1200
 Db 1989 AGATCCGCTACTCCGCGCGGAGGACCTGAGCGAGTCCGCATCGACCGGATCGGAACAC 2048
 QY 1201 TCTCGAGAAAGCGTCTAACCAAGTCACAGTCGCAAGATC 1239
 Db 2049 TCTCGAGAAAGCGTCTAACCAAGTCACAGTCGCAAGATC 2087
 RESULT 7
 ID ADB75132
 XX ADB75132 standard; DNA; 7231 BP.
 XX
 AC ADB75132;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Plasmid pDV80 DNA sequence.
 XX
 KW ophthalmological; antiinflammatory; antidiabetic; gene therapy;
 KW adenovirus inverted terminal repeat sequence;
 KW adenovirus packaging signal; photoreceptor-specific promoter;
 KW adenovirus type 37; adenovirus type D serotype; adenovirus type 2;
 KW adenovirus type 5; photoreceptor; trophic factor; anti-apoptotic factor;
 KW rhodopsin; wild-type Stargardt disease gene; STDG1; anti-cancer agent;
 KW retinal degenerative disease; retinitis pigmentosa; Stargardt's disease;
 KW diabetic retinopathy; retinal vascularisation; choroideraemia;
 KW gyrate atrophy; macular dystrophy; retinoblastoma;
 KW photoreceptor-restricted transgene expression;
 KW recombinant adenovirus vector; adenovirus type 5; plasmid; cyclic;
 KW circular; ds; pDV80.
 XX
 OS Synthetic.
 XX
 PN US2002193327-A1.
 XX
 PD 19-DEC-2002.
 XX
 PF 01-MAY-2001; 2001US-00847101.
 XX
 PR 01-MAY-2000; 2000US-00562934.
 XX
 PA (SCRI) SRIIPPS RES INST.
 XX
 PI Nemerow GR, Von Seggern DJ, Friedlander M;
 XX
 DR WPI; 2003-657234/62.
 XX
 PT Novel nucleic acids comprising adenovirus inverted terminal repeat
 PT sequences, adenovirus packaging signals operatively linked to the
 PT sequences and photoreceptor-specific promoters, useful for treating
 PT retinitis pigmentosa.
 XX
 PS Example 8; Page 99-103; 106pp; English.
 XX
 CC The invention describes an isolated nucleic acid (I) comprising
 CC adenovirus inverted terminal repeat sequence, an adenovirus packaging
 CC signal operatively linked to the sequence, and a photoreceptor-specific
 CC promoter. A Recombinant adenovirus vector (II) comprising (I) is useful
 CC for targeted delivery of a gene product to the eye of a mammal which
 CC involves administering (II) that comprises heterologous DNA encoding the
 CC gene product or resulting in expression of the gene product, where the
 CC recombinant virus comprises a fibre protein that specifically or
 CC selectively binds to receptors that are expressed on cells which are
 CC photoreceptors, in the eye. The recombinant virus comprises a fibre
 CC protein which is an adenovirus type 37, from an adenovirus type D
 CC serotype. The fibre is a chimeric protein containing a sufficient portion
 CC of the N-terminus of an adenovirus type 2 or type 5 fibre protein for

CC interaction with an adenovirus type 2 or type 5 penton, and a sufficient
 CC portion of an adenovirus serotype D knob portion of the fiber for
 CC selective binding to photoreceptors in the eye of a mammal. The
 CC encapsulated nucleic acid comprises a photoreceptor-specific promoter
 CC operatively linked to a nucleic acid comprising the therapeutic product
 CC which is chosen from trophic factor, anti-apoptotic factor, gene encoding
 CC a rhodopsin protein, wild-type Stargardt disease gene (STDG1), an anti-
 CC cancer agent and a protein that regulates expression of a photoreceptor-
 CC specific gene product. The delivery is effected for treatment of an
 CC ocular disease such as retinal degenerative disease e.g., retinitis
 CC pigmentosa, Stargardt's disease, diabetic retinopathies, retinal
 CC vascularisation, choroideraemia, gyrate atrophy or macular dystrophy or
 CC retinoblastoma inherited and acquired retinal and neovascular
 CC degenerative diseases. The viral nucleic acid comprises an adenovirus
 CC inverted terminal repeat (ITR) sequences, and an adenovirus packaging
 CC signal operatively linked to the sequence. The ITRs and packaging signal
 CC are derived from an adenovirus serotype B or C, or adenovirus type 2 or
 CC 5. The viral nucleic acid further comprises a photoreceptor-specific
 CC promoter. (II) includes photoreceptor promoters providing a means not
 CC only for specific targeting of expression in these cells, but also for
 CC photoreceptor-restricted transgene expression. This sequence represents
 CC an adenovirus 37 fibre-expressing plasmid used in the preparation of
 CC adenoviral gene delivery vectors.
 XX
 SQ Sequence 7231 BP; 1751 A; 1807 C; 1869 G; 1804 T; 0 U; 0 Other;
 Query Match 99.9%; Score 1239; DB 10; Length 7231;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGATCCACTCTCTCCGCATCGCTGTCTGCGAGGGCCAGCTGTGGGGTAGTACTCCCT 60
 Db 849 GGATCCACTCTCTCTCCGCATCGCTGTCTGCGAGGGCCAGCTGTGGGGTAGTACTCCCT 908
 QY 61 CTGAAAACGGCGCATGACTTCTCGCTAAGATTCTCAGTTTCCAAAACGAGGAGGATT 120
 Db 909 CTGAAAACGGCGCATGACTTCTCGCTAAGATTCTCAGTTTCCAAAACGAGGAGGATT 968
 QY 121 GATATTCACTGGCCCGCGGTGATGCCCTTTGAGGGTGGCCGCATCCATCTCGTCAGAAAA 180
 Db 969 GATATTCACTGGCCCGCGGTGATGCCCTTTGAGGGTGGCCGCATCCATCTCGTCAGAAAA 1028
 QY 181 GACAATCTTTTGTGTCAAGCTTTGGTGGCAACGACCCGCTAGAGGGGCTTGGACAGCAA 240
 Db 1029 GACAATCTTTTGTGTCAAGCTTTGGTGGCAACGACCCGCTAGAGGGGCTTGGACAGCAA 1088
 QY 241 CTTGCGCATGGAGCGCAGGGTTTGGTTTTCGCGATCGCGCGCTCCTTTGGCGCGCAT 300
 Db 1089 CTTGCGCATGGAGCGCAGGGTTTGGTTTTCGCGATCGCGCGCTCCTTTGGCGCGCAT 1148
 QY 301 GTTTAGCTGCACGTATTTCGCGCGCAACGACCCGCATTCGCGGAAAGACGGTGGTCGCTC 360
 Db 1149 GTTTAGCTGCACGTATTTCGCGCGCAACGACCCGCATTCGCGGAAAGACGGTGGTCGCTC 1208
 QY 361 GTCGGGACACAGGTGACGCGCCAAACCGCGTGTGTGCGGGTGAACAAGTCAACGCTGGT 420
 Db 1209 GTCGGGACACAGGTGACGCGCCAAACCGCGTGTGTGCGGGTGAACAAGTCAACGCTGGT 1268
 QY 421 GGCTTACCTCTCCGCTAGGCGCTGTTGGTCCACAGAGGGCGCGCCCTTTGGCGGACGA 480
 Db 1269 GGCTTACCTCTCCGCTAGGCGCTGTTGGTCCACAGAGGGCGCGCCCTTTGGCGGACGA 1328
 QY 481 GAATGGCGGTAGGGGGTCTAGCTCGCTCTCGTCGGGGGGGTCTGCGTCCACGGTAAAGAC 540
 Db 1329 GAATGGCGGTAGGGGGTCTAGCTCGCTCTCGTCGGGGGGGTCTGCGTCCACGGTAAAGAC 1388
 QY 541 CCCGGGACAGCGCGCGCTGAAGTAGTCTATCTTGCATCTCTTGCATCTAGGCGCTG 600
 Db 1389 CCCGGGACAGCGCGCGCTGAAGTAGTCTATCTTGCATCTCTTGCATCTAGGCGCTG 1448
 QY 601 CTGCGCATCGCGCGCGCGCTCGTATGGTTGAGTGGGGGACCCCAATGGCAT 660
 Db 1449 CTGCGCATCGCGCGCGCGCTCGTATGGTTGAGTGGGGGACCCCAATGGCAT 1508

QY 661 GGGTGGGTAGCGCGGAGCGTATGCGCGCAATGTCGTAAACGTAGAGGGCTCTCT 720
Db |||||
QY 1509 GGGTGGGTAGCGCGGAGCGTATGCGCGCAATGTCGTAAACGTAGAGGGCTCTCT 1568
Db |||||
QY 721 GAGTATCCAGATATGAGGTAGCATCTTCCACCGCGGATGTCGGCGGCGACGTAATC 780
Db |||||
QY 1569 GAGTATCCAGATATGAGGTAGCATCTTCCACCGCGGATGTCGGCGGCGACGTAATC 1628
Db |||||
QY 781 GTATAGTTCTGTCGAGGAGCGAGGAGGTGCGGACCGAGTTGTCACGGCGGCTGCTC 840
Db |||||
QY 1629 GTATAGTTCTGTCGAGGAGCGAGGAGGTGCGGACCGAGTTGTCACGGCGGCTGCTC 1688
Db |||||
QY 841 TGCTCGAAGACTATCTGCTGAAAGATGCGCATGTGAGTTGGATGATATGTTGGACGCTG 900
Db |||||
QY 1689 TGCTCGAAGACTATCTGCTGAAAGATGCGCATGTGAGTTGGATGATATGTTGGACGCTG 1748
Db |||||
QY 901 GAAGAGCTTGAAGCTGGGCTGTGAGACCTACCGGTCACCGCAAGAGGCGTAGGA 960
Db |||||
QY 1749 GAAGAGCTTGAAGCTGGGCTGTGAGACCTACCGGTCACCGCAAGAGGCGTAGGA 1808
Db |||||
QY 961 GTCCGCGAGCTGTTGACAGCTCGCGGTGACCTGACGCTAGGCGGCGAGTTCAG 1020
Db |||||
QY 1809 GTCCGCGAGCTGTTGACAGCTCGCGGTGACCTGACGCTAGGCGGCGAGTTCAG 1868
Db |||||
QY 1021 GGTTCCTTGATGATGTATCTATCTGTCCTCTTTTTCACAGCTCGCGGTTGAG 1080
Db |||||
QY 1869 GGTTCCTTGATGATGTATCTATCTGTCCTCTTTTTCACAGCTCGCGGTTGAG 1928
Db |||||
QY 1081 GACAACTCTTCGCGGTCTTCCAGTACTCTTGGATCGGAACCGGTCGCGCTCGCAAG 1140
Db |||||
QY 1929 GACAACTCTTCGCGGTCTTCCAGTACTCTTGGATCGGAACCGGTCGCGCTCGCAAG 1988
Db |||||
QY 1141 AGATCCGACTCCGCGCGGAGGACCTGAGCGAGTCCGATCGACCGGATCGGAAC 1200
Db |||||
QY 1989 AGATCCGACTCCGCGCGGAGGACCTGAGCGAGTCCGATCGACCGGATCGGAAC 2048
Db |||||
QY 1201 TCTCGAGAAAGCGTCTAACCAAGTACAGTTCGCAAGATC 1239
Db |||||
QY 2049 TCTCGAGAAAGCGTCTAACCAAGTACAGTTCGCAAGATC 2087
Db |||||

RESULT 8

ADP48774
ID ADP48774 standard; DNA; 7231 BP.
XX
AC ADP48774;
XX
DT 12-FEB-2004 (first entry)
XX
DE Adenovirus associated plasmid DNA.
XX
KW cytostatic; anti-HIV; gene therapy; HIV gene expression inhibitor;
KW HIV gene expression activation; adenovirus tripartite leader; TPL;
KW gutless adenoviral vector particle;
KW helper-independent fiberless recombinant adenovirus vector;
KW packaging cell line; pseudotyping; adenovirus vector; gene therapy;
KW hereditary disorder; tumour; HIV infection; ds; cyclic; circular.
XX
OS Synthetic.
OS unidentified adenovirus.
XX
XX
XX US2003157688-A1.
XX
XX 21-AUG-2003.
XX
XX 14-JAN-2000; 2000US-00482682.
XX
XX 14-JAN-1999; 99US-0115920P.
XX 26-JUN-2000; 2000US-00423783.
XX
XX (VSEB/) VON SEGGERN D J.
XX (NEME/) NEMEROW G R.
PA

(HALL/) HALLENBECK P.
PA (STEV/) STEVENSON S.
XX (SKRI/) SKRIPCHENKO Y.
PI Von Seggern DJ, Nemerow GR, Hallenbeck P, Stevenson S;
PI Skripchenko Y;
XX WPI; 2003-843463/78.
XX
PT Novel isolated nucleic acid molecule useful for delivering heterologous
PT gene to human or any animal, or for producing gutless adenoviral vector
PT particle.
XX
PS Claim 10; SEQ ID NO 64; 157pp; English.
XX
CC The invention describes an isolated nucleic acid molecule (I) comprising
CC an adenovirus tripartite leader (TPL) nucleotide, the TPL nucleotide
CC sequence comprising a first and second different TPL exons or first,
CC second and third same or different TPL exons, the TPL exons chosen from
CC complete or partial TPL exon 1, complete TPL exon 2 and complete TPL exon
CC 3. (I) is useful for delivering a heterologous gene to a human or any
CC animal, or for producing a gutless adenoviral vector particle. A
CC recombinant adenovirus particle (II) is useful for delivery of an
CC exogenous gene to a target cell which involves contacting the cell with
CC an amount of (II) sufficient to infect the cell. A helper-independent
CC fiberless recombinant adenovirus vector genome (III) is useful for
CC producing an adenovirus vector particle containing (III) which involves
CC providing a packaging cell line which complements replication and
CC packaging of the genome and (III) which is deficient in expressing
CC sufficient functional fiber protein to support assembly of fiber
CC containing particles and harvesting the particle produced by the cell
CC line. (III) is useful for pseudotyping recombinant viral vectors which
CC involves complementing a missing fiber gene of (III) or helper-dependent
CC fiberless recombinant adenovirus vector genome by expressing in packaging
CC cells a fiber gene from a different adenoviral serotype than the
CC recombinant adenovirus vector. (III) is also useful for specifically
CC targeting an adenovirus vector to a cell of choice. (I) is useful for
CC gene therapy. (II) is useful for treating diseases such as hereditary
CC disorder, and for reducing proliferation of tumour cells in a subject, or
CC to disrupt HIV infection. This sequence represents a plasmid associated
CC with the creation of adenoviral vectors and packaging cell lines.
XX
SQ Sequence 7231 BP; 1751 A; 1807 C; 1869 G; 1804 T; 0 U; 0 Other;

Query Match 99.9%; Score:1239; DB 10; Length 7231;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGATCCACTCTCTTCGCGCATCGTGTCTGCGAGGCGCCAGCTGTGGGGTGAGTACTCCCT 60
Db 849 GGATCCACTCTCTTCGCGCATCGTGTCTGCGAGGCGCCAGCTGTGGGGTGAGTACTCCCT 908
QY 61 CTGAAAGCGGCGCATGACTTCTGGCTAAGATTGTCAAGTTTCCTCAAAAACGAGGAGATT 120
Db 909 CTGAAAGCGGCGCATGACTTCTGGCTAAGATTGTCAAGTTTCCTCAAAAACGAGGAGATT 968
QY 121 GATATTACCTGCGCGCGGTGATGCTCTTTCAGGGTGGCGCATCCATCTGTGTGAGAAA 180
Db 969 GATATTACCTGCGCGCGGTGATGCTCTTTCAGGGTGGCGCATCCATCTGTGTGAGAAA 1028
QY 181 GACAATCTTTTGTGTCAAGCTTGTGGGCAAAACGACCCGTAGAGGGCGTTGGACAGCAA 240
Db 1029 GACAATCTTTTGTGTCAAGCTTGTGGGCAAAACGACCCGTAGAGGGCGTTGGACAGCAA 1088
QY 241 CTTGGCGATGAGCGCAGGGTTGGTTTGTGCGGATCGGCGGCGCTCTTGGCGCGCGAT 300
Db 1089 CTTGGCGATGAGCGCAGGGTTGGTTTGTGCGGATCGGCGGCGCTCTTGGCGCGCGAT 1148
QY 301 GTTTAGCTGCACGTATTTCGCGGCAACGACCGCCATTTCGGGAAGACGGTGGTGGCTC 360
Db 1149 GTTTAGCTGCACGTATTTCGCGGCAACGACCGCCATTTCGGGAAGACGGTGGTGGCTC 1208
QY 361 GTCGGGCACCAGGTGTCACGCGCAACCGCGGTGTGTGAGGGGTGACAAGGTCAACGCTG 420

```
Db 1209 GTCCGGCACCAGTGTCAAGCGCCACCGGGTGTGAGGGTGACAAAGGTCAACGCTGGT 1268
QY 421 GGCTACTCTCCGCTAGCGCTCGTGGTCCAGCAGAGCGCGCCCTTTGCGGAGCA 480
Db 1269 GGTACTCTCTCCGCTAGCGCTCGTGGTCCAGCAGAGCGCGCCCTTTGCGGAGCA 1328
QY 481 GAATGGCGGTAGGGGTCTAGCTGCTCTCGTCCGGGGGGTCTGGTCCACGGTAAAGAC 540
Db 1329 GAATGGCGGTAGGGGTCTAGCTGCTCTCGTCCGGGGGGTCTGGTCCACGGTAAAGAC 1388
QY 541 CCCGGCAGCAGCGCGCTCGAAGTAGTCTATCTTTGCATCTCTTCAAGTCTTAGCGCTG 600
Db 1389 CCCGGCAGCAGCGCGCTCGAAGTAGTCTATCTTTGCATCTCTTCAAGTCTTAGCGCTG 1448
QY 601 CTGCCATCGCGGGCGGCAAGCGCGCTCGTATGGGTTGAGTGGGGACCCCATGGCAT 660
Db 1449 CTGCCATCGCGGGCGGCAAGCGCGCTCGTATGGGTTGAGTGGGGACCCCATGGCAT 1508
QY 661 GGGGTGGGTGAGCGGGAGGCTATCGCGCGCAATGTCCTAAAGTAGAGGGCTCTCT 720
Db 1509 GGGGTGGGTGAGCGGGAGGCTATCGCGCGCAATGTCCTAAAGTAGAGGGCTCTCT 1568
QY 721 GAGTATTCACAGATATGATGGGTAGCATCTTCCACCGCGATGCTGGCGGCACGCTAATC 780
Db 1569 GAGTATTCACAGATATGATGGGTAGCATCTTCCACCGCGATGCTGGCGGCACGCTAATC 1628
QY 781 GTATAGTTCTGTCGAGGAGCGAGAGGTTCGGACCCGAGTTGCTACGGGCGGGTGTCT 840
Db 1629 GTATAGTTCTGTCGAGGAGCGAGAGGTTCGGACCCGAGTTGCTACGGGCGGGTGTCT 1688
QY 841 TGCTCGGAAGACTATCTGCTGAAGATGCGCATGAGTGTGATGATGATGTTGACGCTG 900
Db 1689 TGCTCGGAAGACTATCTGCTGAAGATGCGCATGAGTGTGATGATGATGTTGACGCTG 1748
QY 901 GAAGACGTTGAAGCTGGCGTCTGTGAGACCTACCGCGTCACGCACGAAGGAGCGTAGGA 960
Db 1749 GAAGACGTTGAAGCTGGCGTCTGTGAGACCTACCGCGTCACGCACGAAGGAGCGTAGGA 1808
QY 961 CTCGCGCAGCTTTGTACAGCTCGCGCGTGACCTGCAAGTCTAGGGCGCAGTAGTCCAG 1020
Db 1809 CTCGCGCAGCTTTGTACAGCTCGCGCGTGACCTGCAAGTCTAGGGCGCAGTAGTCCAG 1868
QY 1021 GGTTCCTTGATGATGATCATCTTATCTGTGCTCTTTTTCACAGCTGCGGTTGAG 1080
Db 1869 GGTTCCTTGATGATGATCATCTTATCTGTGCTCTTTTTCACAGCTGCGGTTGAG 1928
QY 1081 GACAAACTCTTCGCGGTCTTTCCAGTACTCTTGGATCGGAAACCGCTCGGAACG 1140
Db 1929 GACAAACTCTTCGCGGTCTTTCCAGTACTCTTGGATCGGAAACCGCTCGGAACG 1988
QY 1141 AGATCCGTACTCCGCGCGGAGGACCTGAGCAGTCCGCAATCGACCGGATCGGAAACG 1200
Db 1989 AGATCCGTACTCCGCGCGGAGGACCTGAGCAGTCCGCAATCGACCGGATCGGAAACG 2048
QY 1201 TCTCAGAAAGCGCTCTAACAGTACAGTCCAGATC 1239
Db 2049 TCTCAGAAAGCGCTCTAACAGTACAGTCCAGATC 2087
```

RESULT 9

AAA59072

ID AAA59072 standard; DNA; 7960 BP.

XX AC

XX AAA59072;

XX 07-NOV-2000 (first entry)

XX Nucleotide sequence of plasmid pdV67.

XX Adenovirus; tripartite leader; adenovirus vector particle; gene delivery;

XX ss.

XX

```
OS Synthetic.
XX WO200042208-A1.
XX 20-JUL-2000.
XX 14-JAN-2000; 2000WO-EP000265.
XX 14-JAN-1999; 99US-0115920P.
XX (NOVS ) NOVARTIS AG.
XX (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.
XX (SCRI ) SCRIPPS RES INST.
XX Nemerow GR, Von Seggern DJ, Hallenbeck PL, Stevenson SC;
XX Skripchenko Y;
XX WPI; 2000-476068/41.
XX New nucleic acid comprising an adenovirus tripartite leader nucleotide
XX for producing high-capacity and targeted vectors for adenovirus-based
XX gene therapy.
XX Claim 10; Page 184-186; 212pp; English.
XX The specification describes a nucleic acid molecule comprising an
XX adenovirus (AV) tripartite leader (TPL) nucleotide with a sequence
XX comprising two different TPL exons or three same or different TPL exons.
XX The nucleic acid is used to produce an adenovirus vector particle,
XX deliver an exogenous gene to a target cell, pseudotype recombinant viral
XX vectors, target an adenovirus vector to a cell, produce a modified
XX adenovirus, deliver a heterologous gene to an animal and produce a
XX gutless adenoviral vector particle. The present sequence represents
XX plasmid pdV67, which contains a TPL
XX
XX Sequence 7960 BP; 1934 A; 2070 C; 1993 G; 1963 T; 0 U; 0 Other;
```

Query March 99.9%; Score 1239; DB 3; Length 7960;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 GGATCCACTCTCTTCCGCATCGCTGTCTGCGAGGCGCAGCTGTGGGGTGAGTACTCCCT 60
Db 929 GGATCCACTCTCTTCCGCATCGCTGTCTGCGAGGCGCAGCTGTGGGGTGAGTACTCCCT 988
QY 61 CTGAAAAAGCGGCGCATGCTTCTGCGCTAAGATTGTCAGTTTCCAAAAACGAGGAGGATTT 120
Db 989 CTGAAAAAGCGGCGCATGCTTCTGCGCTAAGATTGTCAGTTTCCAAAAACGAGGAGGATTT 1048
QY 121 GATATTACCTGGCGCGGCGGATGCTTTGAGGGTGGCGCGCATCTCGTCAGAAAA 180
Db 1049 GATATTACCTGGCGCGGCGGATGCTTTGAGGGTGGCGCGCATCTCGTCAGAAAA 1108
QY 181 GACAACTCTTTTGTGTCAAGCTTTGGTGGCAACGACCCGTAGAGGGCGTTGGACAGCAA 240
Db 1109 GACAACTCTTTTGTGTCAAGCTTTGGTGGCAACGACCCGTAGAGGGCGTTGGACAGCAA 1168
QY 241 CTTGCGCATGAGCGCAGGGGTTTGGTTTTGTCTCGCATCGCGCGCTCTTTGGCGCGCAT 300
Db 1169 CTTGCGCATGAGCGCAGGGGTTTGGTTTTGTCTCGCATCGCGCGCTCTTTGGCGCGCAT 1228
QY 301 GTTTAGCTGCACGTATTTCGCGCGCAACGACCCGCATTCGGGAAAGAGCGTGTGGCTC 360
Db 1229 GTTTAGCTGCACGTATTTCGCGCGCAACGACCCGCATTCGGGAAAGAGCGTGTGGCTC 1288
QY 361 GTCGGGCACAGGTGTCACGCGCCAAACCGCGGTTGTGACGGGTGACAAAGCTCAACGCTG 420
Db 1289 GTCGGGCACAGGTGTCACGCGCCAAACCGCGGTTGTGACGGGTGACAAAGCTCAACGCTG 1348
QY 421 GGCTACTCTCTCCGCTAGGCGCTCGTTGGTCCAGCAGAGGCGCGCGCTTTGGCGGAGCA 480
Db 1349 GGCTACTCTCTCCGCTAGGCGCTCGTTGGTCCAGCAGAGGCGCGCGCTTTGGCGGAGCA 1408
```

QY 481 GAATGCGGTAGGGGCTTAGCTGCTCTCGTCCGGGGGCTCTGCTCAAGTAAAGAC 540
Db 1409 GAATGCGGTAGGGGCTTAGCTGCTCTCGTCCGGGGGCTCTGCTCAAGTAAAGAC 1468
QY 541 CCGGGCAGCAGCGCGCTCGAAGTAGTCTATCTTGCACTCTTGGCATCTTGAAGTCTAGCGCTG 600
Db 1469 CCGGGCAGCAGCGCGCTCGAAGTAGTCTATCTTGGCATCTTGAAGTCTAGCGCTG 1528
QY 601 CTGCCATGCGCGGGCGGCAAGCGCGCTCGTATGGGTTGAGTGGGGGAGCCCATGGCAT 660
Db 1529 CTGCCATGCGCGGGCGGCAAGCGCGCTCGTATGGGTTGAGTGGGGGAGCCCATGGCAT 1588
QY 661 GGGGTGGGTAGCGCGGAGCGGTACATGCGCGCAATGTCTGTAACCTAGAGGGCTCTCT 720
Db 1589 GGGGTGGGTAGCGCGGAGCGGTACATGCGCGCAATGTCTGTAACCTAGAGGGCTCTCT 1648
QY 721 GAGTATTCGAAGATGTAGGTAGCATCTTCCACCGCGGATGCTGCGCGCAGCGTAATC 780
Db 1649 GAGTATTCGAAGATGTAGGTAGCATCTTCCACCGCGGATGCTGCGCGCAGCGTAATC 1708
QY 781 GTATAGTTCTGCGGAGGAGCGAGGAGTGGGACCGAGTTGTACGGGCGGGTGTCTC 840
Db 1709 GTATAGTTCTGCGGAGGAGCGAGGAGTGGGACCGAGTTGTACGGGCGGGTGTCTC 1768
QY 841 TGCTCGGAAGACTATCTGCTGAAAGATGCGCATGTGAGTTGGATGATATGTTGGACGCTG 900
Db 1769 TGCTCGGAAGACTATCTGCTGAAAGATGCGCATGTGAGTTGGATGATATGTTGGACGCTG 1828
QY 901 GAAGACGTTGAGCTGGGCTCTGAGACCTACCGGTCAAGCAGGAGGAGGCGTAGGA 960
Db 1829 GAAGACGTTGAGCTGGGCTCTGAGACCTACCGGTCAAGCAGGAGGAGGCGTAGGA 1888
QY 961 CTCGGCGAGCTTGTGACAGCTCGCGGCTGACCTGACGCTCTAGGGCGCAGTAGTCCAG 1020
Db 1889 CTCGGCGAGCTTGTGACAGCTCGCGGCTGACCTGACGCTCTAGGGCGCAGTAGTCCAG 1948
QY 1021 GGTTCCTTGATGATGATATCTATCTGTCCTCTTTTTCACAGCTCGCGGTGAG 1080
Db 1949 GGTTCCTTGATGATGATATCTATCTGTCCTCTTTTTCACAGCTCGCGGTGAG 2008
QY 1081 GACAACTCTTCGGGCTCTTCCAGTACTCTTGGATCGGAAACCGTCCGACG 1140
Db 2009 GACAACTCTTCGGGCTCTTCCAGTACTCTTGGATCGGAAACCGTCCGACG 2068
QY 1141 AGATCCGTACTCCGCGCGGAGGACCTGAGGAGTCCGATCGACCGGATCGGAAAC 1200
Db 2069 AGATCCGTACTCCGCGCGGAGGACCTGAGGAGTCCGATCGACCGGATCGGAAAC 2128
QY 1201 TCTGAGAAAGCGCTTAACAGTACAGTCTGCAAGATC 1239
Db 2129 TCTGAGAAAGCGCTTAACAGTACAGTCTGCAAGATC 2167

RESULT 10
ID ABA94274
AC ABA94274 standard; DNA; 7960 BP.
XX ABA94274;
XX
DT 13-MAR-2002 (first entry)
XX Nucleotide sequence of adenoviral plasmid pdv67.
DE
XX
KW Adenovirus; inverter terminal repeat sequence; ITRS; ocular disease;
KW fiber protein; photoreceptor; rhodopsin; stargardt disease gene; STDG1;
KW optical; antidiabetic; antidiabetic; cytostatic;
KW gene therapy; tripartite leader; tPL; ss.
XX
OS Synthetic.
XX
FN WO200183729-A2.
XX
PD 08-NOV-2001.

XX
PF
XX 30-APR-2001; 2001WO-EP004863.
XX
PR 01-MAY-2000; 2000US-00562934.
XX
PA (NOVS) NOVARTIS AG.
PA (SCRI) SCRIPPS RES INST.
PA (NEME) NEMEROW G R.
PA (VSEG) VON SEGGERN D J.
PA (FRIE) FRIEDLANDER M.
XX
PI Nemerow GR, Von Seggern DJ, Friedlander M;
XX WPI; 2002-082846/11.
XX
PT Polynucleotide for making vectors, useful for treating ocular diseases,
PT e.g., retinitis pigmentosa, comprises adenovirus inverter terminal repeat
PT sequences, packaging signal and photoreceptor-specific promoter.
XX
PS Example 5; Page 134-136; 149pp; English.
XX
CC The invention provides an isolated polynucleotide comprising adenovirus
CC (AV) inverter terminal repeat sequences (ITRS), AV packaging signal
CC operatively linked to ITRS and a photoreceptor-specific promoter. A
CC recombinant AV vector (AVV) comprising the polynucleotide is useful for
CC targeted delivery of a gene product to the eye (especially to the
CC vitreous cavity), for treating an ocular disease, e.g., retinal
CC degenerative disease, retinitis pigmentosa, Stargardt's disease, diabetic
CC retinopathies, retinal vascularizations, and retinoblastoma, of a mammal
CC preferably human. The AAV comprises a fiber protein that specifically or
CC selectively binds to receptors that are expressed on cells (preferably
CC photoreceptors in the eye). Preferably, the recombinant virus comprise a
CC fiber protein from an adenovirus type D subgroup or is a chimeric protein
CC containing a portion of the N-terminus of an adenovirus type 2 or type 5
CC penton, and the therapeutic product is a trophic factor, an anti-
CC apoptotic factor, gene encoding a rhodopsin protein, a wild-type
CC stargardt disease gene (STDG1), an anti-cancer agent and a protein that
CC regulates expression of a photoreceptor specific gene product. The viral
CC nucleic acid of AAV comprises ITRS and packaging signal derived from AAV
CC subgroup B or C, especially an AV type 2 or type 5. AAV is also useful
CC for targeted gene therapy, where the vector comprises an AV type 37 fiber
CC protein or its portion, and selectively transduces photoreceptors and
CC delivers a gene product encoded by AAV. The present sequence represents
CC the nucleotide sequence of plasmid pdv67, a plasmid containing adenoviral
CC tripartite leader (tPL)
XX
SQ Sequence 7960 BP; 1934 A; 2070 C; 1993 G; 1963 T; 0 U; 0 Other;
Query Match 99.9%; Score 1239; DB 6; Length 7960;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGATCCACTCTCTCCGCATCGTGTCTGCGAGGCCAGCTGTGGGTGAGTACTCCCT 60
Db 929 GGATCCACTCTCTCCGCATCGTGTCTGCGAGGCCAGCTGTGGGTGAGTACTCCCT 988
QY 61 CTGAAAGCGGCGCATGCTCTGCGCTAAGATTGTCTTCCAAAACGAGGAGGATTT 120
Db 989 CTGAAAGCGGCGCATGCTCTGCGCTAAGATTGTCTTCCAAAACGAGGAGGATTT 1048
QY 121 GATATTACCTGGCGCGGCTGATGCCCTTTGAGGGTGGCGCATCCATCTGTCAGAAAA 180
Db 1049 GATATTACCTGGCGCGGCTGATGCCCTTTGAGGGTGGCGCATCCATCTGTCAGAAAA 1108
QY 181 GACAACTCTTTTGTGTCAGCTTGGTGGCAAAACGACCCGTAGAGGGCGTTGGACAGCAA 240
Db 1109 GACAACTCTTTTGTGTCAGCTTGGTGGCAAAACGACCCGTAGAGGGCGTTGGACAGCAA 1168
QY 241 CTTGCGCATGAGCGCAGGGTTGGTTTTTCTCGCATCGCGGCTCTCTTGGCGCGCAT 300
Db 1169 CTTGCGCATGAGCGCAGGGTTGGTTTTTCTCGCATCGCGGCTCTCTTGGCGCGCAT 1228
QY 301 GTTTAGCTGCACGTATTTCGCGCGCAACGACCCGCATTTCGGGAAAGACGGTGGTGGCGGTC 360

```
Db      1229 GTTATGCTGCAGTATTTCGGCGCAACGCAACGCCATTTCGGGAAGACGGTGGTGGCTC 1288
Qy      361 GTGCGGCAACAGGTGCACCGCCAAACCGCGTGTGTGCGAGGTGCAAGGTCAACGCTGGT 420
Db      1289 GTGCGGCAACAGGTGCACCGCCAAACCGCGTGTGTGCGAGGTGCAAGGTCAACGCTGGT 1348
Qy      421 GGCTACCTCTCCGCTAGGCGCTCTGTTGGTCCAGCAGAGGGCGCGCCCTTGC CGCGACCA 480
Db      1349 GGCTACCTCTCCGCTAGGCGCTCTGTTGGTCCAGCAGAGGGCGCGCCCTTGC CGCGACCA 1408
Qy      481 GAATCGCGGTAGGGGCTCTAGCTGCGTCTCGTCCGGGGGCTCTGCGTCCACGCTAAAGAC 540
Db      1409 GAATCGCGGTAGGGGCTCTAGCTGCGTCTCGTCCGGGGGCTCTGCGTCCACGCTAAAGAC 1468
Qy      541 CCCGGGACAGCGCGCGCTGGAAGTAGTCTATCTTTGCATCTTTGCAAGTCTTAGCGCCTG 600
Db      1469 CCGGGGACAGCGCGCGCTGGAAGTAGTCTATCTTTGCATCTTTGCAAGTCTTAGCGCCTG 1528
Qy      601 CTGCCATCGCGGGGGGCAAGCGCGCTCTGATCGGTTGAGTGGGGACCCCATGGCAT 660
Db      1529 CTGCCATCGCGGGGGGCAAGCGCGCTCTGATCGGTTGAGTGGGGACCCCATGGCAT 1588
Qy      661 GGGGTGGGTAGCGCGGCGCTACATCGCGCAAAATGTCGTAACGCTAGAGGGGCTCTCT 720
Db      1589 GGGGTGGGTAGCGCGGCGCTACATCGCGCAAAATGTCGTAACGCTAGAGGGGCTCTCT 1648
Qy      721 GAGTATCCAAAGATATGAGGTAGCATCTTCCACCGCGATGCTGGCGCGCACGTAATC 780
Db      1649 GAGTATCCAAAGATATGAGGTAGCATCTTCCACCGCGATGCTGGCGCGCACGTAATC 1708
Qy      781 GTATAGTTCGTCGAGGAGCGAGAGGTGCGGACCGAGGTGCTACGGGCGGCGTGCCTC 840
Db      1709 GTATAGTTCGTCGAGGAGCGAGAGGTGCGGACCGAGGTGCTACGGGCGGCGTGCCTC 1768
Qy      841 TGCTCGGAAGACTATCTGCTCGAAGATGCGATGTCAGTTGGATGATATGTTGGAAGCTG 900
Db      1769 TGCTCGGAAGACTATCTGCTCGAAGATGCGATGTCAGTTGGATGATATGTTGGAAGCTG 1828
Qy      901 GAAGACGTTGAAGCTGGGCTCTGTGAGACCTACCGCGTCAGCACGAAGAGGCGGTAGGA 960
Db      1829 GAAGACGTTGAAGCTGGGCTCTGTGAGACCTACCGCGTCAGCACGAAGAGGCGGTAGGA 1888
Qy      961 GTGCGGAGCTTGTGACAGCTCGGCGGTGACCTGCGAGTCTAGGGCGCGAGTAGTCCAG 1020
Db      1889 GTGCGGAGCTTGTGACAGCTCGGCGGTGACCTGCGAGTCTAGGGCGCGAGTAGTCCAG 1948
Qy      1021 GGTTCCTTGATGATGTCATCTTATCTGTCCTCTTTTTCACAGCTCGCGGTTGAG 1080
Db      1949 GGTTCCTTGATGATGTCATCTTATCTGTCCTCTTTTTCACAGCTCGCGGTTGAG 2008
Qy      1081 GACAAACTCTTCGCGGCTCTTTCAGTACTCTTGGATCGGAACCCGCTCGGCTCCGAACG 1140
Db      2009 GACAAACTCTTCGCGGCTCTTTCAGTACTCTTGGATCGGAACCCGCTCGGCTCCGAACG 2068
Qy      1141 AGATCCGTACTCGCGCGCGGAGGACCTGAGGAGTCCGCATCGACCGGATCGGAAAC 1200
Db      2069 AGATCCGTACTCGCGCGCGGAGGACCTGAGGAGTCCGCATCGACCGGATCGGAAAC 2128
Qy      1201 TCTCGAGAAAGCGCTTAACCAAGTCACAGTCACAGTCGCAAGATC 1239
Db      2129 TCTCGAGAAAGCGCTTAACCAAGTCACAGTCGCAAGATC 2167
```

RESULT 11

ADB75120

XX ADB75120 standard; DNA; 7960 BP.

AC ADB75120;

XX ADB75120;

DT 04-DEC-2003 (first entry)

XX 04-DEC-2003 (first entry)

DE Plasmid pdv67 DNA sequence.

ophthalmological; antiinflammatory; antidiabetic; gene therapy;
adenovirus inverted terminal repeat sequence;
adenovirus packaging signal; photoreceptor-specific promoter;
adenovirus type 37; adenovirus type D serotype; adenovirus type 2;
adenovirus type 5; photoreceptor; trophic factor; anti-apoptotic factor;
rhodopsin; wild-type Stargardt disease gene; STGD1; anti-cancer agent;
retinal degenerative disease; retinitis pigmentosa; Stargardt's disease;
diabetic retinopathy; retinal vascularisation; choroideraemia;
gyrate atrophy; macular dystrophy; retinoblastoma;
photoreceptor-restricted transgene expression;
recombinant adenovirus vector; adenovirus type 5; plasmid; cyclic;
circular; ds; pdv67; tripartite leader sequence; TPL.
Synthetic.

US2002193327-A1.

19-DEC-2002.

01-MAY-2001; 2001US-00847101.

01-MAY-2000; 2000US-00562934.

(SCRI) SCRIPPS RES INST.

Nemerow GR, Von Seggern DJ, Friedlander M;

WPI; 2003-657234/62.

Novel nucleic acids comprising adenovirus inverted terminal repeat sequences, adenovirus packaging signals operatively linked to the sequences and photoreceptor-specific promoters, useful for treating retinitis pigmentosa.

Example 5; Page 82-86; 106pp; English.

The invention describes an isolated nucleic acid (I) comprising adenovirus inverted terminal repeat sequence, an adenovirus packaging signal operatively linked to the sequence, and a photoreceptor-specific promoter. A Recombinant adenovirus vector (II) comprising (I) is useful for targeted delivery of a gene product to the eye of a mammal which involves administering (II) that comprises heterologous DNA encoding the gene product or resulting in expression of the gene product, where the recombinant virus comprises a fibre protein that specifically or selectively binds to receptors that are expressed on cells which are photoreceptors, in the eye. The recombinant virus comprises a fibre protein which is an adenovirus type 37, from an adenovirus type D serotype. The fibre is a chimeric protein containing a sufficient portion of the N-terminus of an adenovirus type 2 or type 5 fibre protein for interaction with an adenovirus type 2 or type 5 penton, and a sufficient portion of an adenovirus serotype D knob portion of the fibre for selective binding to photoreceptors in the eye of a mammal. The encapsulated nucleic acid comprises a photoreceptor-specific promoter operatively linked to a nucleic acid comprising the therapeutic product which is chosen from tropic factor, anti-apoptotic factor, gene encoding a rhodopsin protein, wild-type Stargardt disease gene (STGD1), an anti-cancer agent and a protein that regulates expression of a photoreceptor-specific gene product. The delivery is effected for treatment of an ocular disease such as retinal degenerative disease e.g., retinitis pigmentosa, Stargardt's disease, diabetic retinopathies, retinal vascularisation, choroideraemia, gyrate atrophy or macular dystrophy or retinoblastoma inherited and acquired retinal and neovascular degenerative diseases. The viral nucleic acid comprises an adenovirus inverted terminal repeat (ITR) sequences, and an adenovirus packaging signal operatively linked to the sequence. The ITRs and packaging signal are derived from an adenovirus serotype B or C, or adenovirus type 2 or 5. The viral nucleic acid further comprises a photoreceptor-specific promoter. (II) includes photoreceptor promoters providing a means not only for specific targeting of expression in these cells, but also for photoreceptor-restricted transgene expression. This sequence represents an adenovirus fibre-expressing plasmid for complementation of fibre-gene-deleted adenoviruses that also comprises the adenovirus tripartite leader

CC sequence for enhancing the expression of complementing adenoviral CC proteins.

XX	Sequence	7960 BP; 1934 A; 2070 C; 1993 G; 1963 T; 0 U; 0 Other;
SQ	Query Match	99.9%; Score 1239; DB 10; Length 7960;
	Best Local Similarity	100.0%; Pred. No. 0;
	Matches 1239; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1	GGATCCACTCTCTTCGCGCATCGCTGTCTGCGAGGCGCAGCTGTTGGGGTGAGTACTCCCT 60
Db	929	GGATCCACTCTCTTCGCGCATCGCTGTCTGCGAGGCGCAGCTGTTGGGGTGAGTACTCCCT 988
Qy	61	CTGAAAAGCGGCGATGACTTCTCGCGCTAAGATTCTCAGTTTCCAAAAACGAGGAGGATTT 120
Db	989	CTGAAAAGCGGCGATGACTTCTCGCGCTAAGATTCTCAGTTTCCAAAAACGAGGAGGATTT 1048
Qy	121	GATATTACCTTGGCCCGCGGTGATGCCCTTTGAGGGTGGCCGCATCCATCTGGTTCAGAAAA 180
Db	1049	GATATTACCTTGGCCCGCGGTGATGCCCTTTGAGGGTGGCCGCATCCATCTGGTTCAGAAAA 1108
Qy	181	GACAACTCTTTTGTGTCAAGCTTGTGGGCAACGACCCTGAGAGGCGCTTGGACAGCAA 240
Db	1109	GACAACTCTTTTGTGTCAAGCTTGTGGGCAACGACCCTGAGAGGCGCTTGGACAGCAA 1168
Qy	241	CTTGGCGATGAGCGCAGGGTTTGCTTTTGTCTCGCGATCGCGCGCTCTCTTGGCGCGCAT 300
Db	1169	CTTGGCGATGAGCGCAGGGTTTGCTTTTGTCTCGCGATCGCGCGCTCTCTTGGCGCGCAT 1228
Qy	301	GTTTAGCTGCAGTATTTCGCGCGCAACGCAACCGCCATTGCGGAAAGACGGTGTGCGGCTC 360
Db	1229	GTTTAGCTGCAGTATTTCGCGCGCAACGCAACCGCCATTGCGGAAAGACGGTGTGCGGCTC 1288
Qy	361	GTCCGGCACCAAGTGCACGCGCCCAACCGCGTGTCTGCAGGGTGCACAGGTCTCAACGCTGGT 420
Db	1289	GTCCGGCACCAAGTGCACGCGCCCAACCGCGTGTCTGCAGGGTGCACAGGTCTCAACGCTGGT 1348
Qy	421	GGCTAACCTCTCCGCGTAGGCGCTCGTTGGTTCACACAGAGGCGGCGCGCTTTGCCCGAGCA 480
Db	1349	GGCTAACCTCTCCGCGTAGGCGCTCGTTGGTTCACACAGAGGCGGCGCGCTTTGCCCGAGCA 1408
Qy	481	GAATGGCGGTAGGGGGTCTAGCTCGCTCTCGTCCGGGGGGTCTGCGTCCACGGTAAAGAC 540
Db	1409	GAATGGCGGTAGGGGGTCTAGCTCGCTCTCGTCCGGGGGGTCTGCGTCCACGGTAAAGAC 1468
Qy	541	CCCCGGCAGCAGGCGCGCTCGAAGTAGTCTATCTTCATCTTGCATCTTGCAGGTCTAGCGGCTG 600
Db	1469	CCCCGGCAGCAGGCGCGCTCGAAGTAGTCTATCTTGCATCTTGCAGGTCTAGCGGCTG 1528
Qy	601	CTGCGCATGCGCGGCGGCAAGCGCGCGCTCGTATGGGTTGAGTGGGGGACCCCATGGCAT 660
Db	1529	CTGCGCATGCGCGGCGGCAAGCGCGCGCTCGTATGGGTTGAGTGGGGGACCCCATGGCAT 1588
Qy	661	GGGTGGGTGAGCCGCGAGGCGGTATACGCCGCAAAATGTCTGTAACGTAGAGGGGCTCTCT 720
Db	1589	GGGTGGGTGAGCCGCGAGGCGGTATACGCCGCAAAATGTCTGTAACGTAGAGGGGCTCTCT 1648
Qy	721	GAGTATTCACAGATATGTAGGGTAGCATCTTCACCGCGGATGCTGGCGCGCAGGTAATC 780
Db	1649	GAGTATTCACAGATATGTAGGGTAGCATCTTCACCGCGGATGCTGGCGCGCAGGTAATC 1708
Qy	781	GTATAGTTTCGTGCGAGGGAGCGAGGAGTCGGGACCGAGGTTGCTACGGCGGGCTGCTC 840
Db	1709	GTATAGTTTCGTGCGAGGGAGCGAGGAGTCGGGACCGAGGTTGCTACGGCGGGCTGCTC 1768
Qy	841	TGCTCGGAAGACTATCTGCCCTGAAAGATGGCATGTGTGAGTTGGATGATATGGTTGGACGCTG 900
Db	1769	TGCTCGGAAGACTATCTGCCCTGAAAGATGGCATGTGTGAGTTGGATGATATGGTTGGACGCTG 1828
Qy	901	GAAAGCGTTGAAGCTGGCGTCTGTGAGACTTACCGCTCACCGCACGAAGGAGGCGTAGGA 960
Db	1829	GAAAGCGTTGAAGCTGGCGTCTGTGAGACTTACCGCTCACCGCACGAAGGAGGCGTAGGA 1888

CC complete or partial TPL exon 1, complete TPL exon 2 and complete TPL exon
CC 3. (I) is useful for delivering a heterologous gene to a human or any
CC animal, or for producing a gutless adenoviral vector particle. A
CC recombinant adenovirus particle (II) is useful for delivery of an
CC exogenous gene to a target cell which involves contacting the cell with
CC an amount of (II) sufficient to infect the cell. A helper-independent
CC fiberless recombinant adenovirus vector genome (III) is useful for
CC producing an adenovirus vector particle containing (III) which involves
CC providing a packaging cell line which complements replication and
CC packaging of the genome and (III) which is deficient in expressing
CC sufficient functional fiber protein to support assembly of fiber
CC containing particles and harvesting the particle produced by the cell
CC line. (III) is useful for pseudotyping recombinant viral vectors which
CC involves complementing a missing fiber gene of (III) or helper-dependent
CC fiberless recombinant adenovirus vector genome by expressing in packaging
CC cells a fiber gene from a different adenoviral serotype than the
CC recombinant adenovirus vector. (III) is also useful for specifically
CC targeting an adenovirus vector to a cell of choice. (I) is useful for
CC gene therapy. (II) is useful for treating diseases such as hereditary
CC disorder, and for reducing proliferation of tumour cells in a subject, or
CC to disrupt HIV infection. This sequence represents an adenovirus
CC tripartite leader sequence added to plasmid pCDN3/fibre to create plasmid
CC pDV67, an adenovirus fibre expressing plasmid for complementation of E4-
CC gene-deleted adenoviruses.

SQ Sequence 7960 BP; 1934 A; 2070 C; 1993 G; 1963 T; 0 U; 0 Other;

Query Match 99.9%; Score 1239; DB 10; Length 7960;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGATCCACTCTTCCGATCGCTCTGCGAGGGCCAGCTGTGGGGTGAGTACTCCCT 60
DB 929 GGATCCACTCTTCCGATCGCTCTGCGAGGGCCAGCTGTGGGGTGAGTACTCCCT 988
QY 61 CTGAAAAGCGGCATGCTTCCGCTAAGATTGTCAGTTTCCAAAACGAGGAGATT 120
DB 989 CTGAAAAGCGGCATGCTTCCGCTAAGATTGTCAGTTTCCAAAACGAGGAGATT 1048
QY 121 GATATTACCTGGCCCGCGGTGATCCCTTTGAGGGTGGCCGATCCATCTGGTCAGAAAA 180
DB 1049 GATATTACCTGGCCCGCGGTGATCCCTTTGAGGGTGGCCGATCCATCTGGTCAGAAAA 1108
QY 181 GACATCTTTTGTGTCAGCTTGGTGGCAACACCGGTAGAGGGGTGGACAGCAA 240
DB 1109 GACATCTTTTGTGTCAGCTTGGTGGCAACACCGGTAGAGGGGTGGACAGCAA 1168
QY 241 CTGGCGATGGAGCGCAGGGTTTGGTTTGTGCGGATCGGCGCTCTTGGCGCGAT 300
DB 1169 CTGGCGATGGAGCGCAGGGTTTGGTTTGTGCGGATCGGCGCTCTTGGCGCGAT 1228
QY 301 GTTTAGCTGCAGTATTCCGCGCAACGACCGCCATTCCGGAAAGACGGTGGTGCCTC 360
DB 1229 GTTTAGCTGCAGTATTCCGCGCAACGACCGCCATTCCGGAAAGACGGTGGTGCCTC 1288
QY 361 GTCGGGCACAGGTGCACCGCCAAACCGGGTTGTGCGAGGTGACAAGTCAACGCTGGT 420
DB 1289 GTCGGGCACAGGTGCACCGCCAAACCGGGTTGTGCGAGGTGACAAGTCAACGCTGGT 1348
QY 421 GGTACCTCTCCGCTAGCGCTCGTGTGTCAGAGAGGGCGGCCCTTTCGCGAGCA 480
DB 1349 GGTACCTCTCCGCTAGCGCTCGTGTGTCAGAGAGGGCGGCCCTTTCGCGAGCA 1408
QY 481 GAATGGCGGTAGGGGTCTAGCTGGCTCTCGTCCGGGGGGTCTGGCTCCACGGTAAAGAC 540
DB 1409 GAATGGCGGTAGGGGTCTAGCTGGCTCTCGTCCGGGGGGTCTGGCTCCACGGTAAAGAC 1468
QY 541 CCCGGCAGCAGCGCGGTGCAAGTGTCTATCTTGCATCTTGCAGTCTAGCGCTG 600
DB 1469 CCCGGCAGCAGCGCGGTGCAAGTGTCTATCTTGCATCTTGCAGTCTAGCGCTG 1528
QY 601 CTGCCATCGCGCGCGCAAGCGCGCTCGTATGGTTGAGTGGGGGACCCCATGGCAT 660
DB |||||

DB 1529 CTGCCATCGCGCGCGCAAGCGCGCTCGTATGGTTGAGTGGGGACCCCATGGCAT 1588
QY 661 GGGTGGGTGAGCGCGAGGCGGTACATGCGCGCAAAATGTCGTAAACGTAGAGGGGTCTCT 720
DB 1589 GGGTGGGTGAGCGCGAGGCGGTACATGCGCGCAAAATGTCGTAAACGTAGAGGGGTCTCT 1648
QY 721 GAGTATTCGAAGATATGATGAGGTAGCATCTTCCACCGGATGCTGGCGCGCAGTAAATC 780
DB 1649 GAGTATTCGAAGATATGATGAGGTAGCATCTTCCACCGGATGCTGGCGCGCAGTAAATC 1708
QY 781 GTATAGTTCGTGCGAGGGAGCGAGGAGTCCGGACCGAGGTTCGTACGGCGGGTGTCTC 840
DB 1709 GTATAGTTCGTGCGAGGGAGCGAGGAGTCCGGACCGAGGTTCGTACGGCGGGTGTCTC 1768
QY 841 TGCTCGGAAGACTATCTGCTCTGAAGATGGCATGTGAGTTGGATGATATGTTGGACGCTG 900
DB 1769 TGCTCGGAAGACTATCTGCTCTGAAGATGGCATGTGAGTTGGATGATATGTTGGACGCTG 1828
QY 901 GAAGACCTTGAAGCTGGCTGTGAGACCTTACCGCTCAGCACGAGGAGGCGTAGGA 960
DB 1829 GAAGACCTTGAAGCTGGCTGTGAGACCTTACCGCTCAGCACGAGGAGGCGTAGGA 1888
QY 961 GTCCGCGAGCTTGTGACCAAGCTCGCGGTGACCTGACACGTCTAGGCGCAGTAGTCCAG 1020
DB 1889 GTCCGCGAGCTTGTGACCAAGCTCGCGGTGACCTGACACGTCTAGGCGCAGTAGTCCAG 1948
QY 1021 GGTTCCTTGATGATGATCATCTTATCTGTGCTCTTTTTCACAGCTCGCGGTGAG 1080
DB 1949 GGTTCCTTGATGATGATCATCTTATCTGTGCTCTTTTTCACAGCTCGCGGTGAG 2008
QY 1081 GACAACTCTTCGCGGTCTTTCAGTACTCTTGGATCGGAAACCGCTCGGCTCCGACG 1140
DB 2009 GACAACTCTTCGCGGTCTTTCAGTACTCTTGGATCGGAAACCGCTCGGCTCCGACG 2068
QY 1141 AGATCCGTACTCCCGCGCGAGGACCTGAGCGAGTCCGATCGACCGGATCGGAAAC 1200
DB 2069 AGATCCGTACTCCCGCGCGAGGACCTGAGCGAGTCCGATCGACCGGATCGGAAAC 2128
QY 1201 TCTCGAAGAGCGCTCTTAAACGATGACAGTCAAGTCCGAGATC 1239
DB 2129 TCTCGAAGAGCGCTCTTAAACGATGACAGTCAAGTCCGAGATC 2167

RESULT 13

AAA59075
ID AAA59075 standard; DNA; 7989 BP.
XX
AC AAA59075;
XX
DT 07-NOV-2000 (first entry)
XX
DE Nucleotide sequence of plasmid pDV69.
XX
KW Adenovirus; tripartite leader; adenovirus vector particle; gene delivery;
XX ss.
XX Synthetic.
XX
XX WO20042208-A1.
XX
XX 20-JUL-2000.
XX
XX 14-JAN-2000; 2000WO-EP000265.
XX
XX 14-JAN-1999; 99US-0115920P.
XX
XX (NOVS) NOVARTIS AG
XX (NOVS) NOVARTIS-ERFINDUNGEN VERW GES MBH.
XX (SCRI) SCRIPPS RES INST.
XX Nemerow GR, Von Seggern DJ, Hallenbeck PL, Stevensen SC;
PI Skripchenko Y;
XX

DR WPI; 2000-476068/41.
XX New nucleic acid comprising an adenovirus tripartite leader nucleotide
PT for producing high-capacity and targeted vectors for adenovirus-based
PT gene therapy.
XX Claim 10; Page 187-190; 212pp; English.
XX The specification describes a nucleic acid molecule comprising an
CC adenovirus (AV) tripartite leader (TPL) nucleotide with a sequence
CC comprising two different TPL exons or three same or different TPL exons.
CC The nucleic acid is used to produce an adenovirus vector particle,
CC deliver an exogenous gene to a target cell, pseudotype recombinant viral
CC vectors, target an adenovirus vector to a cell, produce a modified
CC adenovirus, deliver a heterologous gene to an animal and produce a
CC gutless adenoviral vector particle. The present sequence represents
CC plasmid pdv69, which contains a TPL
XX
SQ Sequence 7989 BP; 1934 A; 2072 C; 1997 G; 1983 T; 0 U; 3 Other;
Query Match 99.9%; Score 1239; DB 3; Length 7989;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGAATCACTCTCTCCGATCGCTGTCTGCGAGGCGCACTTTGGGFTGAGTACTCCCT 60
Db 929 GGAATCACTCTCTCCGATCGCTGTCTGCGAGGCGCACTTTGGGFTGAGTACTCCCT 988
QY 61 CTGAAGAGCGGCATGCTTCTGCGCTAAGATTGTCAGTTTCCAAAACGAGGAGATT 120
Db 989 CTGAAGAGCGGCATGCTTCTGCGCTAAGATTGTCAGTTTCCAAAACGAGGAGATT 1048
QY 121 GATATTACCTGCGCGCGGTCATGCTTTGAGGTTGGCGCATCATCTGTCAGAAAA 180
Db 1049 GATATTACCTGCGCGCGGTCATGCTTTGAGGTTGGCGCATCATCTGTCAGAAAA 1108
QY 181 GACAACTTTTGTGTCAAGCTTGTGTCGCAAAACCGTCAGAGGCGTTGACAGCAA 240
Db 1109 GACAACTTTTGTGTCAAGCTTGTGTCGCAAAACCGTCAGAGGCGTTGACAGCAA 1168
QY 241 CTGCGCATGAGCGCAGGTTTGGTTTGTGCGCATGCGCGCGCTCTTGGCGCGCAT 300
Db 1169 CTGCGCATGAGCGCAGGTTTGGTTTGTGCGCATGCGCGCGCTCTTGGCGCGCAT 1228
QY 301 GTTTAGCTGCAGTATTCGCGCAACGACCGCCATTTCGGAAGAGCGTGTGCGCTC 360
Db 1229 GTTTAGCTGCAGTATTCGCGCAACGACCGCCATTTCGGAAGAGCGTGTGCGCTC 1288
QY 361 GTGCGGCACACAGGTGCACGCGCAACCGCGTTGTGCGAGGTTGACAAGGTCAACGCTG 420
Db 1289 GTGCGGCACACAGGTGCACGCGCAACCGCGTTGTGCGAGGTTGACAAGGTCAACGCTG 1348
QY 421 GGCTACCTCTCCGCTAGCGCTCGTTGTGTCAGAGAGCGCGCGCTTTCGCGGAGCA 480
Db 1349 GGCTACCTCTCCGCTAGCGCTCGTTGTGTCAGAGAGCGCGCGCTTTCGCGGAGCA 1408
QY 481 GAATGCGGTAGGGGTCTAGCTGCTGCTCGTCCGCGGGTCTGCTCCACGTTAAAGAC 540
Db 1409 GAATGCGGTAGGGGTCTAGCTGCTGCTCGTCCGCGGGTCTGCTCCACGTTAAAGAC 1468
QY 541 CCCGGGACAGCGCGCTCGAAGTAGTCTATCTTGCATCTTGCAGTCTTAGCGCTG 600
Db 1469 CCCGGGACAGCGCGCTCGAAGTAGTCTATCTTGCATCTTGCAGTCTTAGCGCTG 1528
QY 601 CTGCCATCGCGCGCGCAAGCGCGCTCGTATGGTTGAGTGGGGGACCCCATGGCAT 660
Db 1529 CTGCCATCGCGCGCGCAAGCGCGCTCGTATGGTTGAGTGGGGGACCCCATGGCAT 1588
QY 661 GGGGTGGGTAGCGCGGAGCGCTACATGCCGCAATGCTTAAAGTAGAGGGGCTCTCT 720
Db 1589 GGGGTGGGTAGCGCGGAGCGCTACATGCCGCAATGCTTAAAGTAGAGGGGCTCTCT 1648
QY 721 GAGTATCCAGATATGATAGGTTAGCATCTTCCACCGCGGATGCTGGCGCGCACGTAATC 780

Db 1649 GAGTATCCAGATATGATAGGTTAGCATCTTCCACCGCGGATGCTGGCGCGCACGTAATC 1708
QY 781 GTATAGTTCTGTCGAGGAGCGAGGAGTCCGGACCGAGGTTGCTACCGGCGGCTGCTC 840
Db 1709 GTATAGTTCTGTCGAGGAGCGAGGAGTCCGGACCGAGGTTGCTACCGGCGGCTGCTC 1768
QY 841 TGCTCGGAAGACTATCTGCTGAAAGATGGCATGTGAGTTGGATGATATGGTTGGACGCTG 900
Db 1769 TGCTCGGAAGACTATCTGCTGAAAGATGGCATGTGAGTTGGATGATATGGTTGGACGCTG 1828
QY 901 GAAGACGTTGAAGCTGGGCTCTGTGAGACCTACCGCTCAGCAGCAAGGAGGCTAGGA 960
Db 1829 GAAGACGTTGAAGCTGGGCTCTGTGAGACCTACCGCTCAGCAGCAAGGAGGCTAGGA 1888
QY 961 GTCCGCGACGCTGTTGACACGCTCGCGGTGACCTGACGCTCTAGGGCGCAGTACTCCAG 1020
Db 1889 GTCCGCGACGCTGTTGACACGCTCGCGGTGACCTGACGCTCTAGGGCGCAGTACTCCAG 1948
QY 1021 GGTTCCTTGATGATGATCATCTTATCTGCTCCCTTTTTCACAGCTCCGCGTTGAG 1080
Db 1949 GGTTCCTTGATGATGATCATCTTATCTGCTCCCTTTTTCACAGCTCCGCGTTGAG 2008
QY 1081 GACAAACTCTTTCGCGGCTCTTTCAGTACTCTTGGATCGGAAACCGCTCGGCTCCGAACG 1140
Db 2009 GACAAACTCTTTCGCGGCTCTTTCAGTACTCTTGGATCGGAAACCGCTCGGCTCCGAACG 2068
QY 1141 AGATCCGCTACTCCGCGCGGAGGACCTGAGCGAGTCCGATCGACCGATCGGAAACCG 1200
Db 2069 AGATCCGCTACTCCGCGCGGAGGACCTGAGCGAGTCCGATCGACCGATCGGAAACCG 2128
QY 1201 TCTCGAGAAAGCGCTCTAAACCAAGTCACAGTCGCAAGATC 1239
Db 2129 TCTCGAGAAAGCGCTCTAAACCAAGTCACAGTCGCAAGATC 2167
RESULT 14
ABA94277
ID ABA94277 standard; DNA; 7989 BP.
XX ABA94277;
XX 13-MAR-2002 (first entry)
XX Nucleotide sequence of adenoviral plasmid pdv69.
XX Adenovirus; inverter terminal repeat sequence; ITRS; ocular disease;
KW fiber protein; photoreceptor; rhodopsin; stargardt disease gene; STDG1;
KW opthalmological; antiinflammatory; antidiabetic; cytostatic;
KW gene therapy; fiber protein; ss.
XX Synthetic.
XX WO200183729-A2.
XX 08-NOV-2001.
XX 30-APR-2001; 2001WO-EP004863.
XX 01-MAY-2000; 2000US-00562934.
XX (NOVS) NOVARTIS AG.
XX (SCRI) SCRIPPS RES INST.
XX (NEME) NEMEROW G R.
XX (VSEG) VON SEGGERN D J.
XX (FRIE) FRIEDLANDER M.
XX Nemerow GR, Von Seggern DJ, Friedlander M;
XX WPI; 2002-082846/11.
XX Polynucleotide for making vectors, useful for treating ocular diseases,
PT e.g., retinitis pigmentosa, comprises adenovirus inverter terminal repeat

PT sequences, packaging signal and photoreceptor-specific promoter.
XX Example 5; Page 137-139; 149pp; English.
PS The invention provides an isolated polynucleotide comprising adenovirus
XX (AV) inverter terminal repeat sequences (ITRS), AV packaging signal
CC operatively linked to ITRS and a photoreceptor-specific promoter. A
CC recombinant AV vector (AV) comprising the polynucleotide is useful for
CC targeted delivery of a gene product to the eye (especially to the
CC vitreous cavity), for treating an ocular disease, e.g., retinal
CC degenerative disease, retinitis pigmentosa, Stargardt's disease, diabetic
CC retinopathies, retinal vascularizations, and retinoblastoma, of a mammal
CC preferably human. The AAV comprises a fiber protein that specifically or
CC selectively binds to receptors that are expressed on cells (preferably
CC photoreceptors in the eye). Preferably, the recombinant virus comprise a
CC fiber protein from an adenovirus type D subgroup or is a chimeric protein
CC containing a portion of the N-terminus of an adenovirus type 2 or type 5
CC penton, and the therapeutic product is a trophic factor, an anti-
CC apoptotic factor, gene encoding a rhodopsin protein, a wild-type
CC stargardt disease gene (STDB1), an anti-cancer agent and a protein that
CC regulates expression of a photoreceptor specific gene product. The viral
CC nucleic acid of AAV comprises ITRS and packaging signal derived from AAV
CC subgroup B or C, especially an AV type 2 or type 5. AAV is also useful
CC for targeted gene therapy, where the vector comprises an AV type 37 fiber
CC protein or its portion, and selectively transduces photoreceptors and
CC delivers a gene product encoded by AAV. The present sequence represents
CC the nucleotide sequence of plasmid pdV69, a plasmid containing a modified
XX adenoviral fiber protein
SQ Sequence 7989 BP; 1934 A; 2072 C; 1997 G; 1983 T; 0 U; 3 Other;
Query Match 99.9%; Score 1239; DB 6; Length 7989;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGATCCACTCTTCGGCATGCTCTCGAGAGCCAGCTGTGGGTGAGTACTCCCT 60
DB 929 GGATCCACTCTTCGGCATGCTCTCGAGAGCCAGCTGTGGGTGAGTACTCCCT 988
QY 61 CTGAAAAGCGGCATGACTTCTCGCTAAGATTGTTCAGTTTCCAAAAACGAGGAGATT 120
DB 989 CTGAAAAGCGGCATGACTTCTCGCTAAGATTGTTCAGTTTCCAAAAACGAGGAGATT 1048
QY 121 GATATTACCTGGCCGGCGGTGATGCTTTGAGGTGGCGCATCCATCTGGTCAGAAAA 180
DB 1049 GATATTACCTGGCCGGCGGTGATGCTTTGAGGTGGCGCATCCATCTGGTCAGAAAA 1108
QY 181 GACAACTTTTGTGTCAGCTTGTGCGCAACACCGGTAGAGGGCGTTGGACAGCAA 240
DB 1109 GACAACTTTTGTGTCAGCTTGTGCGCAACACCGGTAGAGGGCGTTGGACAGCAA 1168
QY 241 CTTGGCGATGAGCGAGCGTTGGTTTGTTCGGCATCGGCGCTCTTGGCGCGCAT 300
DB 1169 CTTGGCGATGAGCGAGCGTTGGTTTGTTCGGCATCGGCGCTCTTGGCGCGCAT 1228
QY 301 GTTTAGCTGCAGTATTTCGGCGCAACCGCACCGCCATTTCGGGAAAGACGGTGGTCGCTC 360
DB 1229 GTTTAGCTGCAGTATTTCGGCGCAACCGCACCGCCATTTCGGGAAAGACGGTGGTCGCTC 1288
QY 361 CTCGGGCAACAGGTGACGCGCAACCGCGTTGTGCAAGGTGACAGGTCAACGCTGGT 420
DB 1289 CTCGGGCAACAGGTGACGCGCAACCGCGTTGTGCAAGGTGACAGGTCAACGCTGGT 1348
QY 421 GCTACCTCTCCCGTAGCGCTCGTTGTGTCAGCAGAGGGCGCGCTTGGCGGAGCA 480
DB 1349 GCTACCTCTCCCGTAGCGCTCGTTGTGTCAGCAGAGGGCGCGCTTGGCGGAGCA 1408
QY 481 GAATCGCGTAGGGGTCTAGCTGCGTCTGTCGTCGGGGGGTCTGCGTCCACGGTTAAGAC 540
DB 1409 GAATCGCGTAGGGGTCTAGCTGCGTCTGTCGTCGGGGGGTCTGCGTCCACGGTTAAGAC 1468
QY 541 CCCGGGACAGCGCGGTGCAAGTGTATCTTTCATCTTCAGAGTCTAGCGCTG 600

DB 1469 CCCGGGACAGCGCGCGCTCGAAGTAGTCTATCTTGCATCTTTCGAAGTCTAGCGCCTG 1528
QY 601 CTGCCATTCGCGCGCGCGCAAGCGCGCTCGTATGGGTTCAGTGGGGAGACCCCATGGCAT 660
DB 1529 CTGCCATTCGCGCGCGCGCAAGCGCGCTCGTATGGGTTCAGTGGGGAGACCCCATGGCAT 1588
QY 661 GGGTGGGTGAGCGCGCGAGGGGTACATGCCCAATCTCGTAAACGTAGAGGGGCTCTCT 720
DB 1589 GGGTGGGTGAGCGCGAGGGGTACATGCCCAATCTCGTAAACGTAGAGGGGCTCTCT 1648
QY 721 GAGTATTTCCAAAGATATGTAGGTAGCATCTTCCACCGCGATGTGGCGCCACACTAATC 780
DB 1649 GAGTATTTCCAAAGATATGTAGGTAGCATCTTCCACCGCGATGTGGCGCCACACTAATC 1708
QY 781 GTATAGTTCGTGCGAGGAGCGAGAGGTGGGACCGAGGTTCGTACGGGCGGCTGCTC 840
DB 1709 GTATAGTTCGTGCGAGGAGCGAGAGGTGGGACCGAGGTTCGTACGGGCGGCTGCTC 1768
QY 841 TGCTCGGAAGACTATCTGCCTGAAGATGGCATGTGAGTTGGATGATATGTTGACGCTG 900
DB 1769 TGCTCGGAAGACTATCTGCCTGAAGATGGCATGTGAGTTGGATGATATGTTGACGCTG 1828
QY 901 GAAGACGTTCAAGCTGGCGTCTGTGACACCTACCGCGTCACGACGAAGAGGCGTAGGA 960
DB 1829 GAAGACGTTCAAGCTGGCGTCTGTGACACCTACCGCGTCACGACGAAGAGGCGTAGGA 1888
QY 961 GTGCGCAGCTTGTTCACAGCTCGGCGGTGACCTGCACTGTTAGGCGGCGAGTAGTCCAG 1020
DB 1889 GTGCGCAGCTTGTTCACAGCTCGGCGGTGACCTGCACTGTTAGGCGGCGAGTAGTCCAG 1948
QY 1021 GGTTCCTTCATGATGTCATATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1080
DB 1949 GGTTCCTTCATGATGTCATATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2008
QY 1081 GACAACTCTTCGCGGTCTTTCCAGTACTCTTTCAGTACGAAACCGTCGGGCTCCGAACG 1140
DB 2009 GACAACTCTTCGCGGTCTTTCCAGTACTCTTTCAGTACGAAACCGTCGGGCTCCGAACG 2068
QY 1141 AGATCCGTACTTCGCGCGCGGAGGACCTGACGCGAGTCCGATCGACCGGATCGGAAAC 1200
DB 2069 AGATCCGTACTTCGCGCGCGGAGGACCTGACGCGAGTCCGATCGACCGGATCGGAAAC 2128
QY 1201 TCTCGAAGAGGCGTCTAACCACTGTCAGTCCGCAAGATC 1239
DB 2129 TCTCGAAGAGGCGTCTAACCACTGTCAGTCCGCAAGATC 2167
RESULT 15
ADB75123
ID ADB75123 standard; DNA; 7989 BP.
XX AC ADB75123;
XX AC
XX AC
DT 04-DEC-2003 (first entry)
XX
XX Plasmid pdV69 DNA sequence.
DE
XX
KW ophthalmological; antiinflammatory; antidiabetic; gene therapy;
KW adenovirus inverted terminal repeat sequence;
KW adenovirus packaging signal; photoreceptor-specific promoter;
KW adenovirus type 37; adenovirus type D serotype; adenovirus type 2;
KW adenovirus type 5; photoreceptor; trophic factor; anti-apoptotic factor;
KW rhodopsin; wild-type Stargardt disease gene; STDB1; anti-cancer agent;
KW retinal degenerative disease; retinitis pigmentosa; Stargardt's disease;
KW diabetic retinopathy; retinal vascularisation; choroideraemia;
KW gyrate atrophy; macular dystrophy; retinoblastoma;
KW photoreceptor-restricted transgene expression;
KW recombinant adenovirus vector; adenovirus type 5; plasmid; cyclic;
KW circular; ds; pdV69; tripartite leader sequence; TPL.
XX OS
XX Synthetic.
XX OS
PN US2002193327-A1.

XX 19-DEC-2002.
 PD
 XX
 PF
 XX
 XX 01-MAY-2001; 2001US-00847101.
 XX
 PR 01-MAY-2000; 2000US-00562934.
 XX
 XX (SRI) SCRIPPS RES INST.
 XX
 XX Nemerow GR, Von Seggern DJ, Friedlander M;
 PI
 XX WPI; 2003-657234/62.
 DR
 XX
 PT Novel nucleic acids comprising adenovirus inverted terminal repeat
 PT sequences, adenovirus packaging signals operatively linked to the
 PT sequences and photoreceptor-specific promoters, useful for treating
 PT retinitis pigmentosa.
 PT
 XX
 PS Example 5; Page 86-90; 106pp; English.
 XX
 CC The invention describes an isolated nucleic acid (I) comprising
 CC adenovirus inverted terminal repeat sequence, an adenovirus packaging
 CC signal operatively linked to the sequence, and a photoreceptor-specific
 CC promoter. A Recombinant adenovirus vector (II) comprising (I) is useful
 CC for targeted delivery of a gene product to the eye of a mammal which
 CC involves administering (II) that comprises heterologous DNA encoding the
 CC gene product or resulting in expression of the gene product, where the
 CC recombinant virus comprises a fibre protein that specifically or
 CC selectively binds to receptors that are expressed on cells which are
 CC photoreceptors, in the eye. The recombinant virus comprises a fibre
 CC protein which is an adenovirus type 37, from an adenovirus type B
 CC serotype. The fibre is a chimeric protein containing a sufficient portion
 CC of the N-terminus of an adenovirus type 2 or type 5 fibre protein for
 CC interaction with an adenovirus type 2 or type 5 penton, and a sufficient
 CC portion of an adenovirus serotype D knob portion of the fiber for
 CC selective binding to photoreceptors in the eye of a mammal. The
 CC encapsulated nucleic acid comprises a photoreceptor-specific promoter
 CC operatively linked to a nucleic acid comprising the therapeutic product
 CC which is chosen from trophic factor, anti-apoptotic factor, gene encoding
 CC a rhodopsin protein, wild-type Stargardt disease gene (STGD1), an anti-
 CC cancer agent and a protein that regulates expression of a photoreceptor-
 CC specific gene product. The delivery is effected for treatment of an
 CC ocular disease such as retinal degenerative disease e.g., retinitis
 CC pigmentosa, Stargardt's disease, diabetic retinopathies, retinal
 CC vascularisation, choroideremia, gyrate atrophy or macular dystrophy or
 CC retinoblastoma inherited and acquired retinal and neovascular
 CC degenerative diseases. The viral nucleic acid comprises an adenovirus
 CC inverted terminal repeat (ITR) sequences, and an adenovirus packaging
 CC signal operatively linked to the sequence. The ITRs and packaging signal
 CC are derived from an adenovirus serotype B or C, or adenovirus type 2 or
 CC 5. The viral nucleic acid further comprises a photoreceptor-specific
 CC promoter. (II) includes photoreceptor promoters providing a means not
 CC only for specific targeting of expression in these cells, but also for
 CC photoreceptor-restricted transgene expression. This sequence represents
 CC an adenovirus fibre-expressing plasmid for complementation of fibre-gene-
 CC deleted adenoviruses that also comprises the adenovirus tripartite leader
 CC sequence for enhancing the expression of complementing adenoviral
 CC proteins.
 XX
 SQ Sequence 7989 BP; 1934 A; 2072 C; 1997 G; 1983 T; 0 U; 3 Other;
 Query Match 99.9%; Score 1239; DB 10; Length 7989;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 GSATCCACTCTCTCCCGCATCGCTCTCGCGAGGCCAGCTGTGGGGTGGAGTCTCCCT 60
 929 GSATCCACTCTCTCCCGCATCGCTCTCGCGAGGCCAGCTGTGGGGTGGAGTCTCCCT 988
 61 CTGAAGAGGGCGCATGCTTCCGCTAAGATTGTCAAGTTTCCTCAAAAACGAGGAGGATT 120
 989 CTGAAGAGGGCGCATGCTTCCGCTAAGATTGTCAAGTTTCCTCAAAAACGAGGAGGATT 1048

QY 121 GATATTCACTGGCCCGCGGTGATGCCCTTTGAGGGTGGCCGCATCCATCTGGTCAGAAAA 180
 DB 1049 GATATTCACTGGCCCGCGGTGATGCCCTTTGAGGGTGGCCGCATCCATCTGGTCAGAAAA 1108
 QY 181 GACAATCTTTTGTGTCAGCTTGTGGCAACGACCCGTAGAGGGGTGGACAGCAA 240
 DB 1109 GACAATCTTTTGTGTCAGCTTGTGGCAACGACCCGTAGAGGGGTGGACAGCAA 1168
 QY 241 CTTGGCGATGAGCGCAGGGTGTGGTTTGTCCGATTCGGCGGGCTCTTTGGCGCGCAT 300
 DB 1169 CTTGGCGATGAGCGCAGGGTGTGGTTTGTCCGATTCGGCGGGCTCTTTGGCGCGCAT 1228
 QY 301 GTTTAGTGCACGTATTCCGCCGCAACGCCACCGCATTCGGGAAAGACGGTGGTGGCTC 360
 DB 1229 GTTTAGTGCACGTATTCCGCCGCAACGCCACCGCATTCGGGAAAGACGGTGGTGGCTC 1288
 QY 361 GTCCGGCACCAGGTGCACGCCCAACCGCGTGTGCAGGGTGACAGGTCAACGCTGGT 420
 DB 1289 GTCCGGCACCAGGTGCACGCCCAACCGCGTGTGTGAGGGTGACAGGTCAACGCTGGT 1348
 QY 421 GGCTACCTCTCCGCGTAGGCGCTCGTTGGTCCAGCAGAGCGCGCCCTTTGCCGAGCA 480
 DB 1349 GGCTACCTCTCCGCGTAGGCGCTCGTTGGTCCAGCAGAGCGCGCCCTTTGCCGAGCA 1408
 QY 481 GAATGGCGGTAGGGGTCTAGCTCGTCTCGTCCGGGGGTCTGCGTCCACGGTAAAGAC 540
 DB 1409 GAATGGCGGTAGGGGTCTAGCTCGTCTCGTCCGGGGGTCTGCGTCCACGGTAAAGAC 1468
 QY 541 CCGGGGACGACGCGCGCGTCCGAGTGTCTATCTTGCATCTTTGCAAGTCTAGGCGCTG 600
 DB 1469 CCGGGGACGACGCGCGCGTCCGAGTGTCTATCTTGCATCTTTGCAAGTCTAGGCGCTG 1528
 QY 601 CTGCGATCGCGCGCGCGCGTCCGAGTGTCTATCTTGCATCTTTGCAAGTCTAGGCGCTG 660
 DB 1529 CTGCGATCGCGCGCGCGCGTCCGAGTGTCTATCTTGCATCTTTGCAAGTCTAGGCGCTG 1588
 QY 661 GGGTGGGTGAGCCGCGAGGCGTACATGCCCAATGTCTGTAAGAGGGGTCTCTCT 720
 DB 1589 GGGTGGGTGAGCCGCGAGGCGTACATGCCCAATGTCTGTAAGAGGGGTCTCTCT 1648
 QY 721 GAGTATTCAGATATGTAGGTTAGTCTTCCACCGCGATGTGGCGCGCACGTATC 780
 DB 1649 GAGTATTCAGATATGTAGGTTAGTCTTCCACCGCGATGTGGCGCGCACGTATC 1708
 QY 781 GTATAGTTCGTGCGAGGAGGAGGAGGTGGGACCGAGGTTCGTACGGCGGGCTGCTC 840
 DB 1709 GTATAGTTCGTGCGAGGAGGAGGAGGTGGGACCGAGGTTCGTACGGCGGGCTGCTC 1768
 QY 841 TGCTCGGAAGACTATCTCCCTGAAGATGGCATGTGAGTTGGATGATATGGTTGACCGCTG 900
 DB 1769 TGCTCGGAAGACTATCTCCCTGAAGATGGCATGTGAGTTGGATGATATGGTTGACCGCTG 1828
 QY 901 GAAGACGTTGAGCTGGCGTCTGAGACCTACCGGTACGCGACGAGAGGGGTAGGA 960
 DB 1829 GAAGACGTTGAGCTGGCGTCTGAGACCTACCGGTACGCGACGAGAGGGGTAGGA 1888
 QY 961 GTCCGCGAGCTTGTTCACGAGTCCGCGGTGACCTGCACGCTTAGGGCGCAGTAGTCCAG 1020
 DB 1889 GTCCGCGAGCTTGTTCACGAGTCCGCGGTGACCTGCACGCTTAGGGCGCAGTAGTCCAG 1948
 QY 1021 GGTTTCCTTGATGATGTCATATCTCTGTCCCTTTTTCACAGCTCGCGGTTGAG 1080
 DB 1949 GGTTTCCTTGATGATGTCATATCTCTGTCCCTTTTTCACAGCTCGCGGTTGAG 2008
 QY 1081 GACAACTCTTTCGCGGTCTTTCAGTACTCTTGGATCGGAAACCCGCTCGGCTCCGAACG 1140
 DB 2009 GACAACTCTTTCGCGGTCTTTCAGTACTCTTGGATCGGAAACCCGCTCGGCTCCGAACG 2068
 QY 1141 AGATCCGTACTCCGCCCGCGAGGACCTGAGCGGAGTCCGATCCAGCGGATCGGAAACC 1200
 DB 2069 AGATCCGTACTCCGCCCGCGAGGACCTGAGCGGAGTCCGATCCAGCGGATCGGAAACC 2128
 QY 1201 TCTCGAAGAGGGCTCTAAACCAAGTCACAGTCGCAAGATC 1239

Db 2129 TCTCGAGAAAGCGTCTTAAACCAAGTCACACAGTCGCAAGATC 2167
|||||

Search completed: July 14, 2005, 07:01:37
Job time : 1751.31 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 05:15:57 ; Search time 11806.6 Seconds
(without alignments)
3997.736 Million cell updates/sec

Title: US-09-482-682-32

Perfect score: 1240
Sequence: 1 ggaaccactctctccgcatt.....cagtcacagtcgcaagatct 1240

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hic:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gsl1:*

9: gb_gsl2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	116.2	9.4	617	9	CL706515
2	115.6	9.3	581	9	CL610425
3	115.6	9.3	637	9	CG899744
4	114.6	9.2	404	9	CL266184
5	108.6	8.8	717	9	CL256302
6	103.6	8.4	728	9	CL256297
7	100.2	8.1	640	9	CL706139
8	88	7.1	579	9	CG691637
9	62.2	5.0	925	9	CNS0091P
10	58.2	4.7	925	9	CNS0091P
11	53.6	4.3	645	9	CC721633
12	52.8	4.3	728	2	BE704468
13	49.2	4.0	645	9	CNS01213
14	48.8	3.9	525	4	BM321004
15	48.2	3.9	516	9	CC624407
16	47.4	3.8	716	4	BG873665
17	47.2	3.8	1012	9	CL476462
18	47	3.8	502	6	CA777137
19	47	3.8	537	5	BO777964
20	47	3.8	3475	3	BC036198
21	46.8	3.8	408	9	CL956771
22	46.8	3.8	604	9	CG049742
23	46.8	3.8	605	9	CG049740
24	46.8	3.8	716	6	CA223022

25	46.8	3.8	889	9	CG344984
26	46.8	3.8	2598	3	AY103647
27	46.6	3.8	463	7	CK122370
28	46.6	3.8	585	4	BI776330
29	46.6	3.8	634	7	CV057862
30	46.6	3.8	704	6	CD935513
31	46.6	3.8	721	7	CV054175
32	46.6	3.8	912	5	BX384097
33	46.6	3.8	942	2	BE214157
34	46.4	3.7	763	4	BI157458
35	46.2	3.7	619	6	CA086410
36	46	3.7	619	9	AG114606
37	46	3.7	798	9	CG433993
38	46	3.7	910	9	CNS0060N
39	46	3.7	1201	9	CNS014BJ
40	45.8	3.7	1133	7	CK209664
41	45.8	3.7	1134	5	BM915656
42	45.8	3.7	1947	9	CL970284
43	45.8	3.7	2332	9	AG363333
44	45.6	3.7	541	6	CD931334
45	45.4	3.7	577	6	CD874810

ALIGNMENTS

RESULT 1

CL706515

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CL706515 617 bp mRNA linear GSS 26-JUL-2004
FHCRC-GT-S22-5F1 2K-GTA Mus musculus cDNA clone FHCRC-GT-S22-5F 5',
mRNA sequence.
CL706515
CL706515.1 GI:50593553
GSS.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Soriano, P.
www.fhcrc.org/labs/soriano/
Unpublished (2003)
Contact: Soriano P
Division of Basic Sciences, A2-025
Fred Hutchinson Cancer Research Center
1100 Fairview Ave. N., Seattle, WA 98109, USA
Tel: 206 667 6825
Fax: 206 667 6522
Email: psoriano@fhcrc.org
ROSAFARY gene trap. The sequence tag is generated by 3'RACE and
represents the 3' insertional cDNA flanking sequence. Additional
information regarding this ES cell line and the insertion mutation
is available upon request at
https://www.fhcrc.org/labs/soriano/otdb/
Class: Gene Trap.
Location/Qualifiers
1. 617
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129S4"
/db_xref="taxon:10090"
/clone="FHCRC-GT-S22-5F"
/sex="Male"
/cell_type="Embryonic stem cell"
/cell_line="AK7.1"
/clone_lib="2K-GTA"
/note="Vector: ROSAFARY"

FEATURES
source

ORIGIN

Query Match 9.4%; Score 116.2; DB 9; Length 617;
Best Local Similarity 84.4%; Pred. No. 1.2e-20;
Matches 130; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 48 GTGAGTACTCCCTCTGAAAGCGGCGCATGACTTCTGCGCTAAGATTGTTCAGTTTCCAAAA 107
 |||||
 Db 1 GTGAGTACTCCCTCTCAAAAGCGGCGCATGACTTCTGCGCTAAGATTGTTCAGTTTCCAAAA 60
 |||||
 QY 108 ACGAGGAGGATTGTATATTCACCTGGCCCGGGTGATGCTTTTGAGGGTGCGCGCATCCA 167
 |||||
 Db 61 ACGAGGAGGATTGTATATTCACCTGGCCCGGGTGATGCTTTTGAGGGTGCGCGCGCA 120
 |||||
 QY 168 TCTGTGTGAGAAAGACAATCTTTTGTGTCAAG 201
 |||||
 Db 121 ACGAAGTTCCTATTCCGAAGTTCCTATTCTCTAG 154
 |||||

RESULT 2
 CL610425 581 bp mRNA linear GSS 01-JUL-2004
 LOCUS
 DEFINITION FHRCR-GT-S17-4D1 2K-GTA Mus musculus cDNA clone FHRCR-GT-S17-4D 5',
 mRNA sequence.

ACCESSION CL610425
 VERSION
 KEYWORDS

SOURCE CL610425.1 GI:48996441

ORGANISM Mus musculus (house mouse)

REFERENCE
 AUTHORS Soriano, P.
 TITLE www.fhrcr.org/labs/soriano/
 JOURNAL Unpublished (2003)
 COMMENT Contact: Soriano P
 Division of Basic Sciences, A2-025
 Fred Hutchinson Cancer Research Center
 1100 Fairview Ave. N., Seattle, WA 98109, USA
 Tel: 206 667 6825
 Fax: 206 667 6522
 Email: psoriano@fhrcr.org

ROSAPARY gene trap. The sequence tag is generated by 3'RACE and
 represents the 3' insertional cDNA flanking sequence. Additional
 information regarding this ES cell line and the insertion mutation
 is available upon request at
 https://www.fhrcr.org/labs/soriano/Gtdb/
 Class: Gene Trap.

FEATURES
 source Location/Qualifiers

1..581
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="129S4"
 /db_xref="taxon:10090"
 /clone="FHRCR-GT-S17-4D"
 /sex="Male"
 /cell_type="Embryonic stem cell"
 /cell_line="AK7.1"
 /clone_lib="2K-GTA"
 /note="Vector: ROSAPARY"

ORIGIN

Query Match 9.3%; Score 115.6; DB 9; Length 581;
 Best Local Similarity 84.4%; Pred. No. 1.8e-20;
 Matches 130; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 48 GTGAGTACTCCCTCTGAAAGCGGCGCATGACTTCTGCGCTAAGATTGTTCAGTTTCCAAAA 107
 |||||
 Db 1 GTGAGTACTCCCTCTCAAAAGCGGCGCATGACTTCTGCGCTAAGATTGTTCAGTTTCCAAAA 60
 |||||
 QY 108 ACGAGGAGGATTGTATATTCACCTGGCCCGGGTGATGCTTTTGAGGGTGCGCGCATCCA 167
 |||||
 Db 61 ACGAGGAGGATTGTATATTCACCTGGCCCGGGTGATGCTTTTGAGGGTGCGCGCGCA 120
 |||||
 QY 168 TCTGTGTGAGAAAGACAATCTTTTGTGTCAAG 201
 |||||
 Db 121 ACGAAGTTCCTATTCCGAAGTTCCTATTCTCTAG 154
 |||||

RESULT 3
 CG899744

LOCUS

DEFINITION FHRCR-GT-S9-7G1 2K-GTA Mus musculus genomic clone FHRCR-GT-S9-7G1
 5', genomic survey sequence.

ACCESSION CG899744

VERSION CG899744.1 GI:39555653

KEYWORDS

SOURCE Mus musculus (house mouse)

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

www.fhrcr.org/labs/soriano/
 Unpublished (2003)
 Contact: Soriano P

Division of Basic Sciences, A2-025

Fred Hutchinson Cancer Research Center

1100 Fairview Ave. N., Seattle, WA 98109, USA

Tel: 206 667 6825

Fax: 206 667 6522

Email: psoriano@fhrcr.org

ROSAPARY gene trap. The sequence tag is generated by 3'RACE and
 represents the 3' insertional cDNA flanking sequence. Additional
 information regarding this ES cell line and the insertion mutation
 is available upon request at
 https://www.fhrcr.org/labs/soriano/Gtdb/
 Class: Gene Trap.

FEATURES
 source Location/Qualifiers

1..637
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="129S4"
 /db_xref="taxon:10090"
 /clone="FHRCR-GT-S9-7G1"
 /sex="Male"
 /cell_type="Embryonic stem cell"
 /cell_line="AK7.1"
 /clone_lib="2K-GTA"
 /note="Vector: ROSAPARY"

ORIGIN

Query Match 9.3%; Score 115.6; DB 9; Length 637;

Best Local Similarity 84.4%; Pred. No. 1.8e-20;

Matches 130; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 48 GTGAGTACTCCCTCTGAAAGCGGCGCATGACTTCTGCGCTAAGATTGTTCAGTTTCCAAAA 107
 |||||

Db 1 GTGAGTACTCCCTCTCAAAAGCGGCGCATGACTTCTGCGCTAAGATTGTTCAGTTTCCAAAA 60
 |||||

QY 108 ACGAGGAGGATTGTATATTCACCTGGCCCGGGTGATGCTTTTGAGGGTGCGCGCATCCA 167
 |||||

Db 61 ACGAGGAGGATTGTATATTCACCTGGCCCGGGTGATGCTTTTGAGGGTGCGCGCGCA 120
 |||||

QY 168 TCTGTGTGAGAAAGACAATCTTTTGTGTCAAG 201
 |||||

Db 121 ACGAAGTTCCTATTCCGAAGTTCCTATTCTCTAG 154
 |||||

RESULT 4

CL266184

LOCUS

DEFINITION FHRCR-GT-S12-1B1 2K-GTA Mus musculus genomic clone FHRCR-GT-S12-1B1
 5', genomic survey sequence.

ACCESSION CL266184

VERSION CL266184.1 GI:42417002

KEYWORDS

SOURCE Mus musculus (house mouse)

ORGANISM

REFERENCE

Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 404)

AUTHORS Soriano, P.
TITLE www.fhcr.org/labs/soriano/
JOURNAL Unpublished (2003)
COMMENT Contact: Soriano P
Division of Basic Sciences, A2-025
Fred Hutchinson Cancer Research Center
1100 Fairview Ave. N., Seattle, WA 98109, USA
Tel: 206 667 6825
Fax: 206 667 6522
Email: psoriano@fhcr.org
ROSAFARY gene trap. The sequence tag is generated by 3'RACE and represents the 3' insertional cDNA flanking sequence. Additional information regarding this ES cell line and the insertion mutation is available upon request at
https://www.fhcr.org/labs/soriano/GTdb/
Class: Gene Trap.

FEATURES source
Location/Qualifiers
1..404
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129S4"
/db_xref="taxon:10090"
/clone="FHCRC-GT-S12-1B1"
/sex="Male"
/cell_type="Embryonic stem cell"
/cell_line="AK7.1"
/clone_lib="2K-GTA"
/note="Vector: ROSAFARY"

ORIGIN
Query Match 9.2%; Score 114.6; DB 9; Length 404;
Best Local Similarity 84.3%; Pred. NO. 3.1e-20;
Matches 129; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 49 TGAGTACTCCCTCTGAAAAGCGGCATGACTTCTCGCTAGATTGTTCAGTTTCCAAAA 108
Db 1 TGAGTACTCCCTCTGAAAAGCGGCATGACTTCTCGCTAGATTGTTCAGTTTCCAAAA 60
QY 109 CGAGGAGATTGATATTACCTGCGCGCGGTGATGCTTTGAGGGTGGCGGCATCCAT 168
Db 61 CGAGGAGATTGATATTACCTGCGCGCGGTGATGCTTTGAGGGTGGCGGCATCCAT 120
QY 169 CTGCTCAGAAAAGACAATCTTTTGTGTCAAG 201
Db 121 CGAAGTCTCTATCCGAGTTCCTATTCTCTAG 153

RESULT 5
CL256302
LOCUS FHCRC-GT-S10-8F1 2K-GTA Mus musculus genomic clone FHCRC-GT-S10-8F1
DEFINITION 5', genomic survey sequence.
ACCESSION CL256302
VERSION CL256302.1 GI:41359955
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 717)
Contact: Soriano P
Division of Basic Sciences, A2-025
Fred Hutchinson Cancer Research Center
1100 Fairview Ave. N., Seattle, WA 98109, USA
Tel: 206 667 6825
Fax: 206 667 6522
Email: psoriano@fhcr.org
ROSAFARY gene trap. The sequence tag is generated by 3'RACE and represents the 3' insertional cDNA flanking sequence. Additional information regarding this ES cell line and the insertion mutation is available upon request at
https://www.fhcr.org/labs/soriano/GTdb/
Class: Gene Trap.

FEATURES source
Location/Qualifiers
1..717
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129S4"
/db_xref="taxon:10090"
/clone="FHCRC-GT-S10-8F1"
/sex="Male"
/cell_type="Embryonic stem cell"
/cell_line="AK7.1"
/clone_lib="2K-GTA"
/note="Vector: ROSAFARY"

ORIGIN
Query Match 8.8%; Score 108.6; DB 9; Length 717;
Best Local Similarity 83.7%; Pred. NO. 1.6e-18;
Matches 123; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 55 CTCCTCTGAAAAGCGGCATGACTTCTCGCTAGATTGTTCAGTTTCCAAAAACGAGA 114
Db 7 CTCCTCTGAAAAGCGGCATGACTTCTCGCTAGATTGTTCAGTTTCCAAAAACGAGA 66
QY 115 GGATTGATATTACCTGCGCGCGGTGATGCTTTGAGGGTGGCGGCATCCATCTGTC 174
Db 67 GGATTGATATTACCTGCGCGCGGTGATGCTTTGAGGGTGGCGGCATCCATCTGTC 126
QY 175 AGAAAAGACAATCTTTTGTGTCAAG 201
Db 127 TCCTATTCGGAAGTTCCTATTCTCTAG 153

RESULT 6
CL256297
LOCUS FHCRC-GT-S10-8A1 2K-GTA Mus musculus genomic clone FHCRC-GT-S10-8A1
DEFINITION 5', genomic survey sequence.
ACCESSION CL256297
VERSION CL256297.1 GI:41359945
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 728)
Contact: Soriano P
Division of Basic Sciences, A2-025
Fred Hutchinson Cancer Research Center
1100 Fairview Ave. N., Seattle, WA 98109, USA
Tel: 206 667 6825
Fax: 206 667 6522
Email: psoriano@fhcr.org
ROSAFARY gene trap. The sequence tag is generated by 3'RACE and represents the 3' insertional cDNA flanking sequence. Additional information regarding this ES cell line and the insertion mutation is available upon request at
https://www.fhcr.org/labs/soriano/GTdb/
Class: Gene Trap.

FEATURES source
Location/Qualifiers
1..728
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129S4"
/db_xref="taxon:10090"
/clone="FHCRC-GT-S10-8A1"
/sex="Male"
/cell_type="Embryonic stem cell"
/cell_line="AK7.1"

is available upon request at
https://www.fhcr.org/labs/soriano/GTdb/
Class: Gene Trap.

FEATURES source
Location/Qualifiers
1..717
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129S4"
/db_xref="taxon:10090"
/clone="FHCRC-GT-S10-8F1"
/sex="Male"
/cell_type="Embryonic stem cell"
/cell_line="AK7.1"
/clone_lib="2K-GTA"
/note="Vector: ROSAFARY"

ORIGIN
Query Match 8.8%; Score 108.6; DB 9; Length 717;
Best Local Similarity 83.7%; Pred. NO. 1.6e-18;
Matches 123; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 55 CTCCTCTGAAAAGCGGCATGACTTCTCGCTAGATTGTTCAGTTTCCAAAAACGAGA 114
Db 7 CTCCTCTGAAAAGCGGCATGACTTCTCGCTAGATTGTTCAGTTTCCAAAAACGAGA 66
QY 115 GGATTGATATTACCTGCGCGCGGTGATGCTTTGAGGGTGGCGGCATCCATCTGTC 174
Db 67 GGATTGATATTACCTGCGCGCGGTGATGCTTTGAGGGTGGCGGCATCCATCTGTC 126
QY 175 AGAAAAGACAATCTTTTGTGTCAAG 201
Db 127 TCCTATTCGGAAGTTCCTATTCTCTAG 153

RESULT 6
CL256297
LOCUS FHCRC-GT-S10-8A1 2K-GTA Mus musculus genomic clone FHCRC-GT-S10-8A1
DEFINITION 5', genomic survey sequence.
ACCESSION CL256297
VERSION CL256297.1 GI:41359945
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 728)
Contact: Soriano P
Division of Basic Sciences, A2-025
Fred Hutchinson Cancer Research Center
1100 Fairview Ave. N., Seattle, WA 98109, USA
Tel: 206 667 6825
Fax: 206 667 6522
Email: psoriano@fhcr.org
ROSAFARY gene trap. The sequence tag is generated by 3'RACE and represents the 3' insertional cDNA flanking sequence. Additional information regarding this ES cell line and the insertion mutation is available upon request at
https://www.fhcr.org/labs/soriano/GTdb/
Class: Gene Trap.

FEATURES source
Location/Qualifiers
1..728
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129S4"
/db_xref="taxon:10090"
/clone="FHCRC-GT-S10-8A1"
/sex="Male"
/cell_type="Embryonic stem cell"
/cell_line="AK7.1"

```

/clone_lib="2K-GTA"
/note="Vector: ROSAFARY"

ORIGIN
Query Match      8.4%; Score 103.6; DB 9; Length 728;
Best Local Similarity 83.8%; Pred. No. 3.7e-17;
Matches 129; Conservative 0; Mismatches 24; Indels 1; Gaps 1;

QY 48 GTGAGTACTCCCTCTGAAAGCGGCATGACTTCTGCGCTAAGATTGTGAGTTTCCAAAA 107
Db 1 GTGAGTACTCCCTCT-CAAAGCGGCATGACTTCTGCGCTAAGATTGTGAGTTTCCAAAA 59

QY 108 ACGAGGAGATTGATATTCACCTGCGCGCGGCTGATGCTTTGAGGTGCGCGCATCCA 167
Db 60 ACGAGGAGATTGATATTCACCTGCGCGCGGCTGATGCTTTGAGGTGCGCGCGCCA 119

QY 168 TCTGCTCAGAAAGACAATCTTTTGTGTCAAG 201
Db 120 ACGAGTTCCTATTCGAGTTCCTATTCCTAG 153

RESULT 7
CL706139      640 bp mRNA linear GSS 20-JUL-2004
LOCUS        FHCRC-GT-S20-7B1 2K-GTA Mus musculus cDNA clone FHCRC-GT-S20-7B 5',
DEFINITION   mRNA sequence.
ACCESSION    CL706139
VERSION      1
KEYWORDS     GSS.
SOURCE       Mus musculus (house mouse)
ORGANISM     Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE    1 (bases 1 to 640)
AUTHORS      Soriano, P.
TITLE        www.fhrc.org/labs/soriano/
JOURNAL      Unpublished (2003)
COMMENT      Contact: Soriano P
              Division of Basic Sciences, A2-025
              Fred Hutchinson Cancer Research Center
              1100 Fairview Ave. N., Seattle, WA 98109, USA
              Tel: 206 667 6825
              Fax: 206 667 6522
              Email: psoriano@fhrc.org
              ROSAFARY gene trap. The sequence tag is generated by 3'RACE and
              represents the 3' insertional cDNA flanking sequence. Additional
              information regarding this ES cell line and the insertion mutation
              is available upon request at
              https://www.fhrc.org/labs/soriano/GTdb/
              Class: Gene Trap.
              Location/Qualifiers
                1..640
                  /organism="Mus musculus"
                  /mol_type="mRNA"
                  /strain="129S4"
                  /db_xref="taxon:10090"
                  /clone="FHCRC-GT-S20-7B"
                  /sex="Male"
                  /cell_type="Embryonic stem cell"
                  /cell_line="AK7.1"
                  /clone_lib="2K-GTA"
                  /note="Vector: ROSAFARY"

FEATURES
source
  Location/Qualifiers
    1..640
      /organism="Mus musculus"
      /mol_type="genomic DNA"
      /strain="129S4"
      /db_xref="taxon:10090"
      /clone="FHCRC-GT-S8-11B"
      /sex="Male"
      /cell_type="Embryonic stem cell"
      /cell_line="AK7.1"
      /clone_lib="2K-GTA"
      /note="Vector: ROSAFARY"

ORIGIN
Query Match      7.1%; Score 88; DB 9; Length 579;
Best Local Similarity 100.0%; Pred. No. 6.9e-13;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 74 ATGACTTCTGCGCTAAGATTGTGAGTTTCCAAAAACGAGGAGATTGATTTACCTGG 133
Db 1 ATGACTTCTGCGCTAAGATTGTGAGTTTCCAAAAACGAGGAGATTGATTTACCTGG 60

QY 134 CCCGCGGTGATGCTTTGAGGTGCGCG 161
Db 61 CCCGCGGTGATGCTTTGAGGTGCGCG 89

RESULT 9
CNS0091P      925 bp DNA linear GSS 03-JUN-1999
LOCUS        Drosophila melanogaster genome survey sequence TEF3 end of BAC #
DEFINITION   BACR19D16 of RPCI-98 library from Drosophila melanogaster (fruit
              fly), genomic survey sequence.
ACCESSION    AL053013
VERSION      AL053013.1 GI:4934461
KEYWORDS     GSS.

/clone_lib="2K-GTA"
/note="Vector: ROSAFARY"

ORIGIN
Query Match      8.1%; Score 100.2; DB 9; Length 640;
Best Local Similarity 83.2%; Pred. No. 3.1e-16;
Matches 114; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 65 AAAGCGGCATGACTTCTGCGCTAAGATTGTGAGTTTCCAAAAACGAGGAGATTGATA 124
Db 16 AAAGCGGCATGACTTCTGCGCTAAGATTGTGAGTTTCCAAAAACGAGGAGATTGATA 75

QY 125 TTCACCTGCGCGCGTATGCTTTGAGGTGCGCGCATCCTGTCAGAAAGACA 184

```

[illegible]

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:39:07 ; Search time 749.127 Seconds
(without alignments)
6468.225 Million cell updates/sec

Title: US-09-482-682-43_COPY_1_100
Perfect score: 100
Sequence: 1 gacggatcgaggatctcccc.....ctgtccctgtgtgtgtgt 100

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*
1: gb_ba.*
2: gb_hgt.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_phi.*
8: gb_pl.*
9: gb_pr.*
10: gb_to.*
11: gb_strs.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100	100.0	3853	6 AR098190	AR098190 Sequence
2	100	100.0	3853	6 AR207832	AR207832 Sequence
3	100	100.0	3853	6 BD009729	BD009729 Tissue sp
4	100	100.0	3886	12 PCDNA3ZEO	X90639 Cloning vec
5	100	100.0	4026	6 AR098191	AR098191 Sequence
6	100	100.0	4026	6 AR207833	AR207833 Sequence
7	100	100.0	4026	6 BD009730	BD009730 Tissue sp
8	100	100.0	4249	6 AR098192	AR098192 Sequence
9	100	100.0	4249	6 AR207834	AR207834 Sequence
10	100	100.0	4249	6 BD009731	BD009731 Tissue sp
11	100	100.0	4341	6 A38214	A38214 Sequence 58
12	100	100.0	4341	6 AX286570	AX286570 Sequence
13	100	100.0	4597	6 AX060344	AX060344 Sequence
14	100	100.0	4840	6 AX133940	AX133940 Sequence
15	100	100.0	5053	6 BD238492	BD238492 Expressio
16	100	100.0	5070	6 AX234391	AX234391 Sequence
17	100	100.0	5082	6 A91754	A91754 Sequence 10
18	100	100.0	5082	6 BD085110	BD085110 Vertebrat
19	100	100.0	5162	6 AX951626	AX951626 Sequence

20	100	100.0	5257	12 CVU89673	US9673 Cloning vec
21	100	100.0	5432	6 BD234590	BD234590 Screening
22	100	100.0	5432	6 AX026821	AX026821 Sequence
23	100	100.0	5446	6 AX319694	AX319694 Sequence
24	100	100.0	5618	6 A44171	A44171 Sequence 1
25	100	100.0	5618	6 AR116416	AR116416 Sequence
26	100	100.0	5618	6 AR222266	AR222266 Sequence
27	100	100.0	5618	6 AR411127	AR411127 Sequence
28	100	100.0	5639	12 AY437643	AY437643 Expressio
29	100	100.0	5651	6 AX211282	AX211282 Sequence
30	100	100.0	5651	6 AX349366	AX349366 Sequence
31	100	100.0	5653	6 I56772	I56772 Sequence 3
32	100	100.0	5653	6 I95540	I95540 Sequence 1
33	100	100.0	5726	12 CVU89672	US9672 Cloning vec
34	100	100.0	5731	6 AX202478	AX202478 Sequence
35	100	100.0	5900	6 AX573107	AX573107 Sequence
36	100	100.0	5995	6 AX685746	AX685746 Sequence
37	100	100.0	6090	6 A63067	A63067 Sequence 11
38	100	100.0	6148	6 BD181637	BD181637 Novel mel
39	100	100.0	6148	6 AX342685	AX342685 Sequence
40	100	100.0	6149	6 BD181638	BD181638 Novel mel
41	100	100.0	6149	6 AX342686	AX342686 Sequence
42	100	100.0	6180	6 AX207724	AX207724 Sequence
43	100	100.0	6186	6 AX211281	AX211281 Sequence
44	100	100.0	6186	6 AX349365	AX349365 Sequence
45	100	100.0	6200	6 BD232461	BD232461 Recombina

ALIGNMENTS

RESULT 1
AR098190
LOCUS AR098190 3853 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 5 from patent US 6074850.
ACCESSION AR098190
VERSION AR098190.1 GI:12807447
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3853)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion polypeptides
JOURNAL Patent: US 6074850-A 5 13-JUN-2000;
FEATURES
source 1..3853
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 100; DB 6; Length 3853;
Best Local Similarity 100.0%; Pred. No. 9.4e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GACGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACATCTGCTCTGATG 60
Db 1 GACGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACATCTGCTCTGATG 60
Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100
RESULT 2
AR207832
LOCUS AR207832 3853 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 5 from patent US 6379927.
ACCESSION AR207832
VERSION AR207832.1 GI:21507688
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

```

Unclassified.
REFERENCE 1 (bases 1 to 3853)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion proteins
JOURNAL Patent: US 6379927-A 5 30-APR-2002;
FEATURES
    source
        1..3853
        /organism="unknown"
        /mol_type="unassigned DNA"

ORIGIN
    Query Match 100.0%; Score 100; DB 6; Length 3853;
    Best Local Similarity 100.0%; Pred. No. 9.4e-24;
    Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGT 100

RESULT 3
BD009729
LOCUS Tissue specific expression of retinoblastoma protein. PAT 31-JAN-2002
DEFINITION Tissue specific expression of retinoblastoma protein.
ACCESSION BD009729
VERSION BD009729.1 GI:18638102
KEYWORDS JP 2001503638-A/3.
SOURCE unidentified
ORGANISM unidentified

REFERENCE 1 (bases 1 to 3853)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Tissue specific expression of retinoblastoma protein
JOURNAL Patent: JP 2001503638-A 3 21-MAR-2001;
COMMENT CANJI INC
    OS Unidentified
    PN JP 2001503638-A/3
    PD 21-MAR-2001
    PF 13-NOV-1997 JP 1998522958
    PR 15-NOV-1996 US 08/751517,14-FEB-1997 US 08/801092 PI
    DOUGLAS ANTELMAN,RICHARD J GREGORY,KENNETH N WILLS PC
    C07H21/04,C07K5/00,A61K38/00,A61K35/12
    CC Strandedness: Single;
    CC Topology: Linear;
    FH Key Location/Qualifiers
    FT CDS 209..862.
    Location/Qualifiers
        1..3853
        /organism="unidentified"
        /mol_type="genomic DNA"
        /db_xref="taxon:32644"

FEATURES
    source
        1..3853
        /organism="unidentified"
        /mol_type="genomic DNA"
        /db_xref="taxon:32644"

ORIGIN
    Query Match 100.0%; Score 100; DB 6; Length 3853;
    Best Local Similarity 100.0%; Pred. No. 9.4e-24;
    Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGT 100

PCDNA3ZEO 3986 bp DNA linear SYN 16-AUG-1995

RESULT 4
PCDNA3ZEO
LOCUS Cloning vector pCDNA3ZEO DNA.
DEFINITION Cloning vector pCDNA3ZEO DNA.
ACCESSION X90639
VERSION X90639.1 GI:9499972
KEYWORDS cloning vector; expression vector; multiple cloning site; Plasmid.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.

REFERENCE 1
AUTHORS Peters,H., Hundhausen,T., Kroenke,M. and Marget,M.
TITLE A new small sized high-level eukaryotic expression vector
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3986)
AUTHORS Peters,H.
TITLE Direct Submission
JOURNAL Submitted (07-AUG-1995) H. Peters, Inst. f. Immunologie,
Michaelistr. 5, D- 24105 Kiel, FRG
COMMENT Related sequences: M21295 and K03104.
FEATURES
    source
        1..3986
        /organism="synthetic construct"
        /mol_type="other DNA"
        /db_xref="taxon:32630"
        /plasmid="pCDNA3ZEO"
        1..2125
        /note="cloning vector (pCDNA3) (Invitrogen)"
        889..994
        /note="multiple cloning site (MCS)"
        2126..2796
        /note="cloning vector (pZeoSV) (Invitrogen)"
        2797..3986
        /note="cloning vector (pCDNA3)"

ORIGIN
    Query Match 100.0%; Score 100; DB 12; Length 3986;
    Best Local Similarity 100.0%; Pred. No. 9.3e-24;
    Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGT 100

RESULT 5
AR098191
LOCUS Sequence 19 from patent US 6074850.
DEFINITION Sequence 19 from patent US 6074850.
ACCESSION AR098191
VERSION AR098191.1 GI:12807448
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 4026)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion polypeptides
JOURNAL Patent: US 6074850-A 19 13-JUN-2000;
FEATURES
    source
        1..4026
        /organism="unknown"
        /mol_type="unassigned DNA"

ORIGIN
    Query Match 100.0%; Score 100; DB 6; Length 4026;
    Best Local Similarity 100.0%; Pred. No. 9.3e-24;
    Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60

```

```
QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100
|||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100

RESULT 6
LOCUS AR207833 4026 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 19 from patent US 6379927.
ACCESSION AR207833
VERSION AR207833.1 GI:21507689
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion proteins
JOURNAL Patent: US 6379927-A 19 30-APR-2002;
FEATURES
source
1. .4026
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN

Query Match 100.0%; Score 100; DB 6; Length 4026;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
|||||
Db 1 GACGGATCGGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100
|||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100

RESULT 8
LOCUS AR098192 4249 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 33 from patent US 6074850.
ACCESSION AR098192
VERSION AR098192.1 GI:12807449
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion polypeptides
JOURNAL Patent: US 6074850-A 33 13-JUN-2000;
FEATURES
source
1. .4249
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN

Query Match 100.0%; Score 100; DB 6; Length 4249;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
|||||
Db 1 GACGGATCGGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100
|||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100

RESULT 9
LOCUS AR207834 4249 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 33 from patent US 6379927.
ACCESSION AR207834
VERSION AR207834.1 GI:21507691
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion proteins
JOURNAL Patent: US 6379927-A 33 30-APR-2002;
FEATURES
source
1. .4249
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN

Query Match 100.0%; Score 100; DB 6; Length 4249;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
|||||
Db 1 GACGGATCGGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100
|||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100

RESULT 7
LOCUS BD009730 4026 bp DNA linear PAT 31-JAN-2002
DEFINITION Tissue specific expression of retinoblastoma protein.
ACCESSION BD009730
VERSION BD009730.1 GI:18638103
KEYWORDS JP 2001503638-A/4.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Tissue specific expression of retinoblastoma protein
JOURNAL Patent: JP 2001503638-A 4 21-MAR-2001;
COMMENT CANJI INC
OS Unidentified
PN JP 2001503638-A/4
PD 21-MAR-2001
PF 13-NOV-1997 JP 1998522958
PR 15-NOV-1996 US 08/751517,14-FEB-1997 US 08/801092 PI
DOUGLAS ANTELMAN,RICHARD J GREGORY,KENNETH N WILLS PC
C07H21/04,C07K5/00,A61K38/00,A61K35/12
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1. .4026
FT /organism='Unidentified'.
FEATURES
source
1. .4026
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN
```

```
Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100

RESULT 10
BD009731
LOCUS BD009731 4249 bp DNA linear PAT 31-JAN-2002
DEFINITION Tissue specific expression of retinoblastoma protein.
ACCESSION BD009731
VERSION BD009731.1 GI:18638104
KEYWORDS JP 2001503638-A/5.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 4249)
AUTHORS Antelman,D., Gregory,R.J., and Wills,K.N.
TITLE Tissue specific expression of retinoblastoma protein
JOURNAL Patent: JP 2001503638-A 5 21-MAR-2001;
CANJ I INC
COMMENT OS Unidentified
PN JP 2001503638-A/5
PD 21-MAR-2001
PF 13-NOV-1997 JP 1998522958
PR 13-NOV-1996 US 08/751517,14-FEB-1997 US 08/801092 PI
DOUGLAS ANTELMAN,RICHARD J GREGORY,KENNETH N WILLS PC
CO7H21/04,C07K5/00,A61K38/00,A61K35/12
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1..4249 /organism='Unidentified'.
FT Location/Qualifiers
1..4249
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4249;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGGTCGACTTCAGTACAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGGTCGACTTCAGTACAATCTGCTCTGATG 60
Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100

RESULT 11
A38214
LOCUS A38214 4341 bp DNA linear PAT 05-MAR-1997
DEFINITION Sequence 58 from Patent WO9408008.
ACCESSION A38214
VERSION A38214.1 GI:2294819
KEYWORDS unidentified
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 4341)
AUTHORS Hawkins,R.E., Russell,S.J., Stevenson,F.K. and Winter,G.P.
TITLE IMPROVEMENTS IN OR RELATING TO IMMUNE RESPONSE MODIFICATION
JOURNAL Patent: WO 9408008-A 58 14-APR-1994;
MEDICAL RES COUNCIL (GB)
OTHER PUBLICATION CA 2145064 940414
OTHER PUBLICATION AU 4832493 940426
OTHER PUBLICATION JP 8501699T 960227.
FEATURES
source 1..4341
Location/Qualifiers

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4249;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGGTCGACTTCAGTACAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGGTCGACTTCAGTACAATCTGCTCTGATG 60
Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100

RESULT 12
AX286570
LOCUS AX286570 4341 bp DNA linear PAT 21-NOV-2001
DEFINITION Sequence 1 from Patent WO0179510.
ACCESSION AX286570
VERSION AX286570.1 GI:17048664
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Rice,J.H. and Stevenson,F.M.
TITLE Materials and methods relating to immune responses to fusion proteins
JOURNAL Patent: WO 0179510-A 1 25-OCT-2001;
Cancer Research Ventures Limited (GB)
FEATURES
source 1..4341
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Vector pVAC1"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4341;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGGTCGACTTCAGTACAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGGTCGACTTCAGTACAATCTGCTCTGATG 60
Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100

RESULT 13
AX060344
LOCUS AX060344 4597 bp DNA linear PAT 22-JAN-2001
DEFINITION Sequence 3 from Patent WO0078358.
ACCESSION AX060344
VERSION AX060344.1 GI:12405832
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Chen,W.
TITLE Hyaluronic acid microspheres for sustained gene transfer
JOURNAL Patent: WO 0078358-A 3 28-DEC-2000;
The Collaborative Group, Ltd. (US)
FEATURES
source 1..4597
Location/Qualifiers
/organism="synthetic construct"
```

```
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="pCDNA3.1/GS vector by Invitrogen Corporation"

ORIGIN
Query Match      100.0%; Score 100; DB 6; Length 4597;
Best Local Similarity 100.0%; Pred. No. 9.2e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
    |||||||
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
    |||||||

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
    |||||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
    |||||||

RESULT 14
AX133940
LOCUS      AX133940      4840 bp      DNA      linear      PAT 15-MAY-2001
DEFINITION Sequence 1 from Patent WO0119853.
ACCESSION AX133940
VERSION   AX133940.1 GI:14139881
KEYWORDS
SOURCE    synthetic construct
ORGANISM  other sequences; artificial sequences.
REFERENCE
AUTHORS   Hollander,A.P., Barker,M.D. and Kafienah,W.C.
TITLE     Cell transfection
JOURNAL   Patent: WO 0119853-A 1 22-MAR-2001;
          THE UNIVERSITY OF SHEFFIELD (GB)
FEATURES
source    1..4840
          /organism="synthetic construct"
          /mol_type="unassigned DNA"
          /db_xref="taxon:32630"
          /note="This sequence is artificial and is based on well
          established commercially available vectors that are cited
          with their vendor within the patent application"

ORIGIN
Query Match      100.0%; Score 100; DB 6; Length 4840;
Best Local Similarity 100.0%; Pred. No. 9.2e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
    |||||||
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
    |||||||

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
    |||||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
    |||||||

RESULT 15
BD238492
LOCUS      BD238492      5053 bp      DNA      linear      PAT 17-JUL-2003
DEFINITION Expression vectors for stimulating an immune response and methods
          of using the same.
ACCESSION BD238492
VERSION   BD238492.1 GI:33048262
KEYWORDS  JP 2002520000-A/18.
SOURCE    synthetic construct
ORGANISM  other sequences; artificial sequences.
REFERENCE
AUTHORS   Fikes,J.D., Hermanson,G.G., Sette,A., Ishioka,G.Y., Livingston,B.
          and Chesnut,R.W.
TITLE     Expression vectors for stimulating an immune response and methods
          of using the same
JOURNAL   Patent: JP 2002520000-A 18 09-JUL-2002;
```

```
EPIMMUNE INC
OS Artificial Sequence
PN JP 2002520000-A/18
PD 09-JUL-2002
PF 13-MAY-1999 JP 2000548449
PR 13-MAY-1998 US 09/078904,15-MAY-1998 US 60/085751 PI
JOHN D FIKES,GARY G HERMANSON,ALESSANDRO
SETTE,GLENN Y ISHIOKA,
PI BRIAN LIVINGSTON,ROBERT W CHESNUT
PC C12N15/09,A61K31/711,A61K39/00,A61K39/12,A61K39/21,A61K39/29,
A61P31/14,
PC A61P31/16,A61P31/20,A61P37/02,C12N15/00
CC vector pEP2
FH Key Location/Qualifiers
FT source 1..5053
          /organism="Artificial Sequence".
          Location/Qualifiers
          1..5053
          /organism="synthetic construct"
          /mol_type="genomic DNA"
          /db_xref="taxon:32630"

ORIGIN
Query Match      100.0%; Score 100; DB 6; Length 5053;
Best Local Similarity 100.0%; Pred. No. 9.2e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
    |||||||
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
    |||||||

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
    |||||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
    |||||||

Search completed: July 14, 2005, 14:03:29
Job time : 749.127 secs
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:35:42 ; Search time 140.988 Seconds
(without alignments)
4198.742 Million cell updates/sec

Title: US-09-482-682-43_COPY_1_100
Perfect score: 100
Sequence: 1 gacgatcgagatctcccc.....ctgtccctgtgtgtgtt 100

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	100	100.0	1506	12 ADM41035	Adm41035 Fungus nu
2	100	100.0	1600	2 ADH11349	Adh11349 Vertebrat
3	100	100.0	1782	12 ADM41037	Adm41037 Cytoegal
4	100	100.0	2241	12 ADM41034	Adm41034 Human nuc
5	100	100.0	2294	12 ADM41036	Adm41036 Cytoegal
6	100	100.0	3853	2 AAV40006	AAV40006 Plasmid p
7	100	100.0	4026	2 AAV40007	AAV40007 Plasmid p
8	100	100.0	4249	2 AAV63466	AAV63466 Plasmid p
9	100	100.0	4341	2 AAG62391	AAG62391 Vector pV
10	100	100.0	4341	6 AAS17704	Aas17704 Vector pV
11	100	100.0	4341	6 ABN83143	Abn83143 Plasmid p
12	100	100.0	4597	4 AAF24901	AAf24901 Nucleotid
13	100	100.0	4639	6 AAD39652	AAd39652 Human sma
14	100	100.0	4840	4 AAF83146	AAf83146 Complete
15	100	100.0	5015	10 ADB33528	Adb33528 Expressio
16	100	100.0	5053	3 AAZ38633	Aaz38633 pEP2 expr
17	100	100.0	5070	4 AAS12839	AAs12839 DNA seque
18	100	100.0	5082	2 ADH11417	Adh11417 Plasmid p
19	100	100.0	5162	10 ADF10526	Adf10526 Plasmid p
20	100	100.0	5162	10 ACC44637	Acc44637 Murine rD

21	100	100.0	5172	13 ADS75099	Ads75099 Plasmid p
22	100	100.0	5192	10 ACC44692	Acc44692 Plasmid p
23	100	100.0	5271	10 ABV77540	Abv77540 Plasmid p
24	100	100.0	5283	10 ABV77538	Abv77538 Plasmid p
25	100	100.0	5293	10 ABV77549	Abv77549 Plasmid p
26	100	100.0	5302	12 ADI34681	Adi34681 Nucleotid
27	100	100.0	5304	10 ABV77539	Abv77539 Plasmid p
28	100	100.0	5425	2 ADH11233	Adh11233 Vertebrat
29	100	100.0	5431	6 ABN86685	Abn86685 Nucleotid
30	100	100.0	5431	10 ADE21866	Ade21866 Plasmid v
31	100	100.0	5431	12 ADO05277	Ado05277 pcDNA3 pl
32	100	100.0	5432	3 AAZ89476	Aaz89476 Transgeni
33	100	100.0	5446	6 AAS18619	AAs18619 Renilla l
34	100	100.0	5446	6 ABL53540	AbL53540 Vector pc
35	100	100.0	5446	12 ADN36314	Adn36314 Plasmid p
36	100	100.0	5458	6 ABL58494	AbL58494 Recombina
37	100	100.0	5458	6 ABL58493	AbL58493 Recombina
38	100	100.0	5543	6 ABK88868	AbK88868 Topoisome
39	100	100.0	5543	12 ADE83791	Ade83791 Plasmid p
40	100	100.0	5543	12 ADO06720	Ado06720 Recombina
41	100	100.0	5614	6 ABL58489	AbL58489 Recombina
42	100	100.0	5614	6 ABL58490	AbL58490 Recombina
43	100	100.0	5618	2 AAQ88310	Aaq88310 Plasmid p
44	100	100.0	5651	5 AAI66195	Aai66195 Human FSH
45	100	100.0	5651	6 ABK40237	AbK40237 DNA encod

ALIGNMENTS.

RESULT 1
ADM41035
ID ADM41035 standard; DNA; 1506 Bp.
XX
AC ADM41035;
XX
DT 17-JUN-2004 (first entry)
XX
DE Fungus nucleotide sequence SEQ ID NO:3.

KW engrafting foreign replacement cell; implanting foreign replacement cell;
KW growth; differentiation; drug development; vaccine development;
KW tissue transplantation; human disease study; fungus; gene; ds.

OS Unidentified.
PN WO2004027029-A2.
XX
PD 01-APR-2004.

PF 17-SEP-2003; 2003WO-US029251.
PR 19-SEP-2002; 2002US-0411790P.

XX (XIME-) XIMEREX INC.

XX Beschornier WE, Sosa CE, Thompson SC;

XX WPI; 2004-295402/27.

XX Engrafting foreign replacement cells within a fetal non-human mammal,
XX useful in producing chimeric mammals, comprises selectively destroying
XX native cells in a tissue of a fetal non-human mammal host.

XX Disclosure; SEQ ID NO 3; 48pp; English.

XX The present invention describes a method for engrafting foreign
XX replacement cells within a foetal non-human mammal, which comprises
XX selectively destroying native cells in a tissue of a foetal non-human
XX mammal host, where the number of maternal cells of the same tissue is not
XX substantially reduced, and implanting foreign replacement cells in the
XX tissue of the fetal non-human mammal host, where the foreign replacement
XX cells replace destroyed cells of the tissue. The method is useful for

CC facilitating growth and differentiation of foreign cells within a
 CC mammalian host, and for producing chimeric mammals that can be used to
 CC develop new drugs and vaccine, factors, drugs and tissues for
 CC transplantation, also useful to study human diseases. The present
 CC sequence represents a nucleotide sequence given in the Sequence Listing
 CC of the present invention but not mentioned further within the
 CC specification.

XX
 SQ Sequence 1506 BP; 454 A; 277 C; 361 G; 414 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 12; Length 1506;
 Best Local Similarity 100.0%; Pred. No. 4e-26;
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
 Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
 Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

RESULT 2
 ADH11349
 ID ADH11349 standard; DNA; 1600 BP.

XX
 AC ADH11349;
 DT 11-MAR-2004 (first entry)

XX
 DE Vertebrate UNC-53 protein homologue related nucleotide sequence.

XX
 KW UNC-53 vertebrate protein homologue; UNC-53; Caenorhabditis elegans;
 KW cell shape regulator; cell motility regulator; cell migration;
 KW cell behaviour regulator; phenotype; signal transduction pathway;
 KW signal transducing protein; signal integrator protein;
 KW neuronal regeneration; revascularisation; wound healing;
 KW chronic neurodegenerative disease; acute traumatic injury;
 KW fibrotic disease; gene; ds.

XX
 OS Unidentified.

XX
 PN WO9824810-A2.

XX
 PD 11-JUN-1998.

XX
 PF 03-DEC-1997; 97WO-EP006956.

XX
 PR 04-DEC-1996; 96GB-00025283.

XX
 (JANC) JANSSEN PHARM NV.

XX
 PI Platteeuw CJ, Buesa Arjol CM, Deraeymaeker M, Verhasselt P;
 PI Pujol NJR, Maertens LJS, Luyten W, Geerts H, Vandekerckhove JS;
 PI Geysen J, Bogaert TAOE;

XX
 WPI; 1998-362411/31.

DR
 P-PSDB; ADH11350.

XX
 PT Vertebrate homologues of C. elegans UNC-53 protein - useful for, e.g.
 PT promoting neuronal regeneration, treating chronic neuro-degenerative
 PT diseases or acute traumatic injuries.

XX
 PS Disclosure; Page 410-411; 479pp; English.

XX
 CC The present invention describes a vertebrate protein homologue of an UNC-
 CC 53 protein of Caenorhabditis elegans or a functional equivalent,
 CC derivative or bioprecursor of UNC-53. Also described: (1) a cDNA sequence
 CC encoding a vertebrate homologue of the C. elegans UNC-53 protein; (2) a
 CC nucleic acid which hybridises to the cDNA of (1); (3) vector comprising
 CC the cDNA as in (1); (4) a host cell containing the vector as in (3); (5)
 CC a transgenic cell, tissue or animal comprising the vector as in (3); (6)

CC a compound identified as an enhancer or inhibitor of the regulation of
 CC cell shape, motility, or the direction of cell migration for use as a
 CC therapeutic; (7) a method for determination of whether a protein is an
 CC inhibitor or enhancer of regulation of cell behaviour, growth, shape or
 CC motility or the direction of migration by contacting a host cell
 CC expressing a homologue of UNC-53 and determining a change of phenotype;
 CC (8) a method for identification of vertebrate homologues of C. elegans
 CC unc-53 gene by hybridising a probe of 15-50 bp of the unc-53 sequence to
 CC a DNA library; and (9) a method for identification of a protein which is
 CC active in the signal transduction pathway of a cell of which a vertebrate
 CC homologue of UNC-53 is a component comprising: (i) contacting an extract
 CC of a cell with an antibody to the UNC-53 homologue; (ii) identifying an
 CC antibody/homologue complex; and (iii) analysing such a complex to
 CC identify any non-antibody protein bound to the complex. UNC-53 is a
 CC signal transducing or signal integrator protein involved in controlling
 CC directionality of cell migration and cell shape in C. elegans. Vertebrate
 CC homologues of UNC-53 can be used to promote neuronal regeneration,
 CC revascularisation or wound healing, to treat chronic neurodegenerative
 CC diseases or acute traumatic injuries or fibrotic diseases. The present
 CC sequence is used in the exemplification of the present invention.

XX
 SQ Sequence 1600 BP; 397 A; 409 C; 398 G; 396 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 2; Length 1600;
 Best Local Similarity 100.0%; Pred. No. 4.1e-26;
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
 Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
 Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

RESULT 3
 ADM41037
 ID ADM41037 standard; DNA; 1782 BP.

XX
 AC ADM41037;
 DT 17-JUN-2004 (first entry)

XX
 DE Cytomegalovirus nucleotide sequence SEQ ID NO:5.

XX
 KW engrafting foreign replacement cell; implanting foreign replacement cell;
 KW growth; differentiation; drug development; vaccine development;
 KW tissue transplantation; human disease study; cytomegalovirus; gene; ds.

XX
 OS Cytomegalovirus.

XX
 PN WO2004027029-A2.

XX
 PD 01-APR-2004.

XX
 PF 17-SEP-2003; 2003WO-US029251.

XX
 PR 19-SEP-2002; 2002US-0411790P.

XX
 (XIME-) XIMEREX INC.

XX
 PI Beschorner WE, Sosa CE, Thompson SC;

XX
 WPI; 2004-295402/27.

XX
 PT Engrafting foreign replacement cells within a fetal non-human mammal,
 PT useful in producing chimeric mammals, comprises selectively destroying
 PT native cells in a tissue of a fetal non-human mammal host.

XX
 PS Disclosure; SEQ ID NO 5; 48pp; English.

XX
 CC The present invention describes a method for engrafting foreign

CC replacement cells within a foetal non-human mammal, which comprises
CC selectively destroying native cells in a tissue of a foetal non-human
CC mammal host, where the number of maternal cells of the same tissue is not
CC substantially reduced, and implanting foreign replacement cells in the
CC tissue of the foetal non-human mammal host, where the foreign replacement
CC cells replace destroyed cells of the tissue. The method is useful for
CC facilitating growth and differentiation of foreign cells within a
CC mammalian host, and for producing chimeric mammals that can be used to
CC develop new drugs and vaccine, factors, drugs and tissues for
CC transplantation, also useful to study human diseases. The present
CC sequence represents a nucleotide sequence given in the Sequence Listing
CC of the present invention but not mentioned further within the
CC specification.

XX SQ Sequence 1782 BP; 458 A; 399 C; 451 G; 474 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 12; Length 1782;
Best Local Similarity 100.0%; Pred. No. 4.2e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCGATCCCTATGCTCGACTCTCAGTACAATCTGCTTGATG 60
Db 1 GACGGATCGGAGATCTCCGATCCCTATGCTCGACTCTCAGTACAATCTGCTTGATG 60

QY 61 CGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100

Db 61 CGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100

RESULT 4

ADM41034

ID ADM41034 standard; DNA; 2241 BP.

XX AC ADM41034;

XX DT 17-JUN-2004 (first entry)

XX DE Human nucleotide sequence SEQ ID NO:2.

XX KW engrafting foreign replacement cell; implanting foreign replacement cell;
XX growth; differentiation; drug development; vaccine development;
XX KW tissue transplantation; human disease study; human; gene; ds.

XX OS Homo sapiens.

XX PN WO2004027029-A2.

XX PD 01-APR-2004.

XX PF 17-SEP-2003; 2003WO-US029251.

XX PR 19-SEP-2002; 2002US-0411790P.

XX PA (XIME-) XIMEREX INC.

XX PI Beschornier WE, Sosa CE, Thompson SC;

XX DR WPI; 2004-295402/27.

XX KW Engrafting foreign replacement cells within a foetal non-human mammal,
XX useful in producing chimeric mammals, comprises selectively destroying
XX native cells in a tissue of a foetal non-human mammal host.

XX PS Disclosure; SEQ ID NO 2; 48pp; English.

XX CC The present invention describes a method for engrafting foreign
CC replacement cells within a foetal non-human mammal, which comprises
CC selectively destroying native cells in a tissue of a foetal non-human
CC mammal host, where the number of maternal cells of the same tissue is not
CC substantially reduced, and implanting foreign replacement cells in the
CC tissue of the foetal non-human mammal host, where the foreign replacement
CC cells replace destroyed cells of the tissue. The method is useful for
CC facilitating growth and differentiation of foreign cells within a

CC mammalian host, and for producing chimeric mammals that can be used to
CC develop new drugs and vaccine, factors, drugs and tissues for
CC transplantation, also useful to study human diseases. The present
CC sequence represents a nucleotide sequence given in the Sequence Listing
CC of the present invention but not mentioned further within the
CC specification.

XX SQ Sequence 2241 BP; 498 A; 630 C; 609 G; 504 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 12; Length 2241;
Best Local Similarity 100.0%; Pred. No. 4.5e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCGATCCCTATGCTCGACTCTCAGTACAATCTGCTTGATG 60
Db 1 GACGGATCGGAGATCTCCGATCCCTATGCTCGACTCTCAGTACAATCTGCTTGATG 60

QY 61 CGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100

Db 61 CGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100

RESULT 5

ADM41036

ID ADM41036 standard; DNA; 2294 BP.

XX AC ADM41036;

XX DT 17-JUN-2004 (first entry)

XX DE Cytomegalovirus nucleotide sequence SEQ ID NO:4.

XX KW engrafting foreign replacement cell; implanting foreign replacement cell;
XX growth; differentiation; drug development; vaccine development;
XX KW tissue transplantation; human disease study; cytomegalovirus; gene; ds.

XX OS Cytomegalovirus.

XX PN WO2004027029-A2.

XX PD 01-APR-2004.

XX PF 17-SEP-2003; 2003WO-US029251.

XX PR 19-SEP-2002; 2002US-0411790P.

XX PA (XIME-) XIMEREX INC.

XX PI Beschornier WE, Sosa CE, Thompson SC;

XX DR WPI; 2004-295402/27.

XX KW Engrafting foreign replacement cells within a foetal non-human mammal,
XX useful in producing chimeric mammals, comprises selectively destroying
XX native cells in a tissue of a foetal non-human mammal host.

XX PS Disclosure; SEQ ID NO 4; 48pp; English.

XX CC The present invention describes a method for engrafting foreign
CC replacement cells within a foetal non-human mammal, which comprises
CC selectively destroying native cells in a tissue of a foetal non-human
CC mammal host, where the number of maternal cells of the same tissue is not
CC substantially reduced, and implanting foreign replacement cells in the
CC tissue of the foetal non-human mammal host, where the foreign replacement
CC cells replace destroyed cells of the tissue. The method is useful for
CC facilitating growth and differentiation of foreign cells within a
CC mammalian host, and for producing chimeric mammals that can be used to
CC develop new drugs and vaccine, factors, drugs and tissues for
CC transplantation, also useful to study human diseases. The present
CC sequence represents a nucleotide sequence given in the Sequence Listing
CC of the present invention but not mentioned further within the
CC specification.

```

SQ Sequence 2294 BP; 466 A; 680 C; 641 G; 507 T; 0 U; 0 Other;

Query Match      100.0%; Score 100; DB 12; Length 2294;
Best Local Similarity 100.0%; Pred. No. 4.5e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTTCAGTACAAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTTCAGTACAAATCTGCTCTGATG 60

Qy 61 CGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100
Db 61 CGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100

RESULT 6
AAV40006
ID AAV40006 standard; DNA; 3853 BP.
AC AAV40006;
XX
XX
DT 27-AUG-2003 (revised)
DT 15-FEB-1999 (first entry)
DE Plasmid pCTM.
XX
KW E2F; transcription factor; human; retinoblastoma protein RB;
KW bladder cancer; restenosis; angioplasty; diabetic retinopathy;
KW thyroid hyperplasia; Grave's disease; psoriasis;
KW benign prostatic hypertrophy; Li-Fraumeni syndrome;
KW peripheral vascular disease; therapy; plasmid pCTM; ss.
XX
OS Human cytomegalovirus.
OS mastadenovirus.
OS unidentified bacteriophage; T7.
OS unidentified bacteriophage; SP6.
OS Macaca mulatta; polyoma virus.
OS Bos taurus.
OS Chimeric.
XX
FH Key
FH promoter
FT 209..864
FT /tag= a
FT /note= "CMV promoter"
FT misc_feature
FT 907..1131
FT /tag= b
FT /function= "tripartite leader sequence"
FT promoter
FT 1132..1149
FT /tag= c
FT /note= "SP6 promoter"
FT misc_feature
FT 1679..3853
FT /tag= d
FT /note= "pUC19 backbone H3 to AatII"
FT CDS
FT complement(2857..3717)
FT /tag= e
FT /note= "AMP-ORF"
XX
XX
XX WO9821228-A1.
XX
XX
XX 22-MAY-1998.
XX
XX 13-NOV-1997; 97WO-US021821.
XX
XX 15-NOV-1996; 96US-00751517.
XX 14-FEB-1997; 97US-00801092.
XX
XX (CANJ-) CANJ INC.
XX
XX Antelman D, Gregory RJ, Wills KN;
XX
XX WPI; 1998-297858/26.
XX
XX New fusion polypeptide of, e.g. transcription factor - used to treat,

```

```

PT e.g. hyper-proliferative disease such as cancer and restenosis.
XX
XX Example 1; Fig 4; 9lpp; English.
PS This is the nucleotide sequence of pCTM, a plasmid which contains a CMV
CC promoter, a tripartite adenovirus leader flanked by T7 and SP6 promoters,
CC and a multiple cloning site with a bovine growth hormone polyA site and
CC downstream SV40 polyA site. It has been used as a vector for the
CC expression of fusion proteins of the invention that comprise
CC retinoblastoma protein (BP, see AAW62465) and E2F transcription factor
CC (see AAW62464). Such fusion proteins, particularly expressed from gene
CC therapy vectors, are used to treat hyperproliferative conditions,
CC specifically cancer (particularly of the bladder) or restenosis. They are
CC more effective in repressing transcription of the E2F promoter than RB
CC alone and cause cell-cycle arrest in a variety of cells. (Updated on 27-
CC AUG-2003 to correct OS field.)
XX
SQ Sequence 3853 BP; 936 A; 986 C; 942 G; 989 T; 0 U; 0 Other;

Query Match      100.0%; Score 100; DB 2; Length 3853;
Best Local Similarity 100.0%; Pred. No. 5.2e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTTCAGTACAAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTTCAGTACAAATCTGCTCTGATG 60

Qy 61 CGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100
Db 61 CGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100

RESULT 7
AAV40007
ID AAV40007 standard; DNA; 4026 BP.
XX
XX
AC AAV40007;
XX
XX
DT 27-AUG-2003 (revised)
DT 15-FEB-1999 (first entry)
DE Plasmid pCTMI.
XX
XX
XX E2F; transcription factor; human; retinoblastoma protein RB;
XX bladder cancer; restenosis; angioplasty; diabetic retinopathy;
XX thyroid hyperplasia; Grave's disease; psoriasis;
XX benign prostatic hypertrophy; Li-Fraumeni syndrome;
XX peripheral vascular disease; therapy; plasmid pCTMI; ss.
XX
XX Human cytomegalovirus.
XX mastadenovirus.
XX unidentified bacteriophage; T7.
XX unidentified bacteriophage; SP6.
XX Macaca mulatta; polyoma virus.
XX Bos taurus.
XX Chimeric.
XX
XX Key
XX promoter
FT 209..864
FT /tag= a
FT /note= "CMV promoter"
FT misc_feature
FT 907..1074
FT /tag= b
FT /function= "tripartite leader sequence"
FT intron
FT 1075..1253
FT /tag= c
FT /note= "hybrid SV40 late intron"
FT promoter
FT 1305..1322
FT /tag= d
FT /note= "SP6 promoter"
FT misc_feature
FT 1851..4026
FT /tag= e
FT /note= "pUC19 backbone H3 to AatII"

```

```
FT CDS complement(3032..3890)
FT FT /*tag= f
XX OS Chimeric.
XX PN
XX XX
XX PD
XX XX
XX PF 13-NOV-1997; 97WO-US021821.
XX PR 15-NOV-1996; 96US-00751517.
XX PR 14-FEB-1997; 97US-00801092.
XX XX
XX PA (CANJ-) CANJI INC.
XX PI Antelman D, Gregory RJ, Wills KN;
XX XX
XX DR WPI; 1998-297858/26.
XX XX
XX FT New fusion polypeptide of, e.g. transcription factor - used to treat,
XX FT e.g. hyper-proliferative disease such as cancer and restenosis.
XX PS Example 1; Fig 6; 91pp; English.
XX CC This is the nucleotide sequence of pCTMI, a plasmid that was constructed
XX CC from pCTM (see AAV40006) by digesting pCTM with XhoI and NotI and
XX CC subcloning a 180 bp intron XhoI-NotI fragment from a pCMV-beta-gal
XX CC vector. Plasmid pCTMI has been used as a vector for the expression of
XX CC fusion proteins of the invention that comprise retinoblastoma protein
XX CC (BP, see AAW62465) and E2F transcription factor (see AAW62464). Such
XX CC fusion proteins, particularly expressed from gene therapy vectors, are
XX CC used to treat hyperproliferative conditions, specifically cancer
XX CC (particularly of the bladder) or restenosis. They are more effective in
XX CC repressing transcription of the E2F promoter than RB alone and cause cell
XX CC -cycle arrest in a variety of cells. (Updated on 27-AUG-2003 to correct
XX CC OS field.)
XX SQ Sequence 4026 BP; 978 A; 1021 C; 982 G; 1045 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 2; Length 4026;
Best Local Similarity 100.0%; Pred. No. 5.3e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGATCGGAGATCTCCGATCCCGATGTCGACTCTCAGTACAATCTGCTGTGATG 60
Dd |||||
Dd 1 GACGATCGGAGATCTCCGATCCCGATGTCGACTCTCAGTACAATCTGCTGTGATG 60
QY 61 CGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
Dd |||||
Dd 61 CGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

RESULT 8
AAV63466
ID AAV63466 standard; DNA; 4249 BP.
XX AC AAV63466;
XX XX
XX DT 27-AUG-2003 (revised)
XX DT 15-FEB-1999 (first entry)
XX XX
XX DE Plasmid pCTMIE.
XX XX
XX KW E2F; transcription factor; human; retinoblastoma protein RB;
XX KW bladder cancer; restenosis; angioplasty; diabetic retinopathy;
XX KW thyroid hyperplasia; Grave's disease; psoriasis;
XX KW benign prostatic hypertrophy; Li-Fraumeni syndrome;
XX KW peripheral vascular disease; therapy; plasmid pCTMIE; ss.
XX OS Human cytomegalovirus.
OS mastadenovirus.
OS unidentified bacteriophage; T7.
OS unidentified bacteriophage; SP6.
```

```
OS Macaca mulatta; polyoma virus.
OS Bos taurus.
XX OS Chimeric.
XX FH Key
XX FT Promoter
XX FT 209..864
XX FT /*tag= a
XX FT /note= "CMV promoter"
XX FT 907..1074
XX FT /*tag= b
XX FT /function= "tripartite leader sequence"
XX FT 1081..1145
XX FT /*tag= c
XX FT /note= "hybrid SV40 late intron"
XX FT 1164..1366
XX FT /*tag= d
XX FT /note= "early mRNA"
XX FT 1261..1332
XX FT /*tag= e
XX FT /note= "72 bp tandem repeat enhancer"
XX FT 1333..1404
XX FT /*tag= f
XX FT /note= "72 bp tandem repeat enhancer"
XX FT 1366
XX FT misc_binding
XX FT /*tag= g
XX FT /note= "T antigen binding site"
XX FT 1372..1478
XX FT /*tag= h
XX FT /note= "hybrid SV40 late intron"
XX FT 1530..1545
XX FT /*tag= i
XX FT /note= "SP6 promoter"
XX FT 2075..4249
XX FT /*tag= j
XX FT /note= "pUC19 backbone H3 to AatII"
XX FT complement(3255..4113)
XX FT /*tag= k
XX FT /note= "AMP-ORF"
XX XX
XX PN WO9821228-A1.
XX XX
XX PD 22-MAY-1998.
XX XX
XX PF 13-NOV-1997; 97WO-US021821.
XX PR 15-NOV-1996; 96US-00751517.
XX PR 14-FEB-1997; 97US-00801092.
XX XX
XX PA (CANJ-) CANJI INC.
XX XX
XX PI Antelman D, Gregory RJ, Wills KN;
XX XX
XX DR WPI; 1998-297858/26.
XX XX
XX FT New fusion polypeptide of, e.g. transcription factor - used to treat,
XX FT e.g. hyper-proliferative disease such as cancer and restenosis.
XX PS Example 1; Fig 8; 91pp; English.
XX CC This is the nucleotide sequence of pCTMIE, a plasmid that was constructed
XX CC by amplifying the SV40 enhancer from SV40 viral DNA by PCR, digesting the
XX CC amplified product with BglII and inserting into BamHI-digested plasmid
XX CC pCTMI (see AAV40007). Plasmid pCTMIE has been used as a vector for the
XX CC expression of fusion proteins of the invention that comprise
XX CC retinoblastoma protein (BP, see AAW62465) and E2F transcription factor
XX CC (see AAW62464). Such fusion proteins, particularly expressed from gene
XX CC therapy vectors, are used to treat hyperproliferative conditions,
XX CC specifically cancer (particularly of the bladder) or restenosis. They are
XX CC more effective in repressing transcription of the E2F promoter than RB
XX CC alone and cause cell-cycle arrest in a variety of cells. (Updated on 27-
XX CC AUG-2003 to correct OS field.)
XX SQ Sequence 4249 BP; 1020 A; 1074 C; 1048 G; 1107 T; 0 U; 0 Other;
```

Query Match 100.0%; Score 100; DB 2; Length 4249;
 Best Local Similarity 100.0%; Pred. No. 5.4e-26;
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
 |||||
 Db 1 GACGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
 |||||

QY 61 CGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100
 |||||
 Db 61 CGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100
 |||||

RESULT 9
 AAQ62391
 ID AAQ62391 standard; DNA; 4341 BP.
 XX
 AC AAQ62391;
 DE Vector pVAC1.
 KW Vector; pVAC1; pRC/RSV; leader sequence; termination signal;
 KW fusion protein; pSfi/NotI; pE1B leader; human; immunoglobulin; VH1;
 KW single chain; Fv; murine antibody; retroviral; envelope; plasmid;
 KW vaccine; ss.
 XX
 OS Synthetic.

FH Key Location/Qualifiers
 FT misc_RNA complement(1. .775)
 FT /tag= c
 FT /note= "Claim 9"
 FT misc_RNA 606. .780
 FT /tag= b
 FT /note= "Claim 8"
 FT misc_RNA 606. .716
 FT /tag= a
 FT /note= "Claim 7"
 FT
 FT
 XX WO9408008-A1.
 XX
 PD 14-APR-1994.
 XX
 PF 04-OCT-1993; 93WO-GB002054.
 XX
 PR 02-OCT-1992; 92GB-00020808.
 XX
 PA (MEDI-) MEDICAL RES COUNCIL.
 XX
 PI Hawkins RE, Russell SJ, Stevenson FK, Winter GP;
 XX
 DR WPI; 1994-135575/16.
 XX
 PT Modulating immune response to a disease marker - by administering a
 PT vector which expresses the disease marker to interact with the immune
 PT system.
 XX
 PS Claim 10; Fig 7; 77pp; English.
 XX
 CC This sequence represents the vector pVAC1. This vector is based on the
 CC commercially available vector pRC/RSV. Leader sequences and termination
 CC signals were introduced into the vector to allow for production of fusion
 CC proteins. The vector, pSfi/NotI, was modified to replace the pE1B
 CC leader with the human immunoglobulin VH1 leader sequence that permits the
 CC encoding of an SfiI cloning site without modification of the amino acid
 CC sequence. This fragment was then cloned as an EcoRI/Blunt-HindIII
 CC fragment into NotI/Blunt-HindIII cut vector pRC/RSV to give pVAC1. The
 CC single chain Fv for an individual patient can be inserted within the VH1
 CC leader sequence. This plasmid when encoding a single chain murine

CC antibody/retroviral envelope fusion protein can be used as a plasmid
 CC vaccine and it induces a strong humoral response to the antibody moiety
 CC in BALB/c mice. (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 4341 BP; 1032 A; 1099 C; 1091 G; 1119 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 2; Length 4341;
 Best Local Similarity 100.0%; Pred. No. 5.4e-26;
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
 |||||
 Db 1 GACGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
 |||||

QY 61 CGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100
 |||||
 Db 61 CGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100
 |||||

RESULT 10
 AAS17704
 ID AAS17704 standard; DNA; 4341 BP.
 XX
 AC AAS17704;
 DT 12-MAR-2002 (first entry)
 XX
 DE Vector pVAC1 encoding a DNA vaccine.
 XX
 KW Cytostatic; vaccine; tetanus toxin; FrC; tumour; CTL; PCR primer; pVAC1;
 KW ds.
 XX
 OS Clostridium tetani.
 OS Homo sapiens.
 OS Synthetic.
 OS Cauliflower mosaic virus.
 XX
 PN WO200179510-A1.
 XX
 PD 25-OCT-2001.
 XX
 PF 17-APR-2001; 2001WO-GB001719.
 XX
 PR 17-APR-2000; 2000GB-00009470.
 XX
 PA (CANC-) CANCER RES VENTURES LTD.
 XX
 PI Rice J, Stevenson F;
 XX
 DR WPI; 2002-066370/09.
 XX
 PT Nucleic acid construct, useful to immunize against various diseases
 PT including cancer, expresses the first domain of tetanus toxin FrC fused
 PT to a disease peptide antigen to provide a vaccine.
 XX
 PS Disclosure; Fig 4; 71pp; English.
 XX
 CC The invention relates to a nucleic acid construct for delivery into
 CC living cells in vivo, to induce an immune response to a disease peptide
 CC antigen, where the construct directs expression of a fusion protein
 CC comprising the peptide antigen and the first domain of FrC. Also included
 CC are a nucleic acid vector comprising the above construct, a host cell
 CC comprising the above construct or vector and a method of producing a
 CC nucleic acid construct for inducing an immune response. The method
 CC comprises identifying a nucleic acid sequence encoding a disease peptide
 CC antigen comprising epitopes characteristic of the disease, cloning the
 CC nucleic acid sequence, introducing the cloned nucleic acid into a vector
 CC which allows the antigen to be expressed as a fusion with a first domain
 CC of FrC from tetanus toxin, and optionally isolating the construct from the
 CC vector. The construct or vector is used as a vaccine to induce an immune
 CC response, particularly to tumour antigens. The present sequence is vector
 CC pVAC1 which encodes a vaccine of the invention
 XX

SQ Sequence 4341 BP; 1033 A; 1099 C; 1090 G; 1119 T; 0 U; 0 Other;
Query Match 100.0%; Score 100; DB 6; Length 4341;
Best Local Similarity 100.0%; Pred. No. 5.4e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGGAGATCTCCGATCCCTATGCTCGACTCTCAGTACAAATCTGCTCTGATG 60
DB 1 GACGGATCGGGAGATCTCCGATCCCTATGCTCGACTCTCAGTACAAATCTGCTCTGATG 60

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100
DB 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100

RESULT 11
AEN831143
ID AEN831143 standard; DNA; 4341 BP.
XX
AC AEN831143;
XX
DT 10-SEP-2002 (first entry)
DE Plasmid pVAC1 complete sequence.
XX
KW Immune response; plant viral coat protein; pVAC1; cytostatic; virucide;
KW cancer; B cell malignancy; ds.
XX
OS Synthetic.
XX
PN WO200240513-A2.
XX
PD 23-MAY-2002.
XX
PF 20-NOV-2001; 2001WO-GB005142.
XX
PR 20-NOV-2000; 2000GB-00028319.
XX
PA (CANC-) 'CANCER RES VENTURES LTD.
PI Savelyeva N, Stevenson F;
XX
XX WPI; 2002-500202/53.
XX
PS Nucleic acid construct for delivery into living cells as a vaccine,
PT useful for treating e.g. cancer, directs the expression of a fusion
PT protein comprising an antigen and an adjuvant sequence derived from a
PT plant viral coat protein.
XX
XX
XX Example 3; Fig 7; 84pp; English.
XX
CC The invention relates to a novel nucleic acid construct for inducing an
CC immune response in vivo to an antigen, capable of directing the
CC expression of a fusion protein that comprises an antigen and an adjuvant
CC sequence derived from a plant viral coat protein. The construct of the
CC invention has cytostatic and virucide activity. The nucleic acid
CC construct is useful for inducing an immune response in a patient, for
CC vaccinating a patient against an infectious disease caused by an antigen
CC derived from a pathogen e.g. a virus, for treating a cancer patient or a
CC patient with a predisposition to cancer and for treating a patient having
CC a B cell malignancy, where the construct is encapsidated, and optionally,
CC a second nucleic acid sequence encoding a further immunomodulatory
CC polypeptide is administered to the patient. The construct is also useful
CC in medical treatment, and in the preparation of a vaccine for treating or
CC preventing a disease state associated with the antigen. The sequence
CC shows the complete sequence of vector pVAC1
XX
SQ Sequence 4341 BP; 1033 A; 1099 C; 1090 G; 1119 T; 0 U; 0 Other;
Query Match 100.0%; Score 100; DB 6; Length 4341;
Best Local Similarity 100.0%; Pred. No. 5.4e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGGAGATCTCCGATCCCTATGCTCGACTCTCAGTACAAATCTGCTCTGATG 60
DB 1 GACGGATCGGGAGATCTCCGATCCCTATGCTCGACTCTCAGTACAAATCTGCTCTGATG 60

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100
DB 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100

RESULT 12
AAF24901
ID AAF24901 standard; DNA; 4597 BP.
XX
AC AAF24901;
XX
DT 20-APR-2001 (first entry)
DE Nucleotide sequence of the plasmid pCDNA3.1/GS.
XX
KW Microspheré; dihydrazide; hyaluronic acid; inflammatory response;
KW myocardial ischemia; cardiac angiogenesis; haemophilia;
KW vascular endothelial growth factor; VEGF; ss.
XX
OS Synthetic.
XX
PN WO200078358-A2.
XX
PD 28-DEC-2000.
XX
PF 19-JUN-2000; 2000WO-US016837.
XX
PR 18-JUN-1999; 99US-0140260P.
XX
PA (COLL-) COLLABORATIVE GROUP LTD.
XX
PI Chen W;
XX
DR WPI; 2001-071363/08.
XX
PT Hyaluronic acid micro spheres for use in gene therapy of myocardial
PT ischemia and hemophilia, comprising dihydrazide derivatized hyaluronic
PT acids crosslinked to nucleic acids.
XX
XX Example 1; Page 36-38; 38pp; English.
PS The specification describes a microsphere comprising dihydrazide
CC derivatised hyaluronic acid crosslinked to a nucleic acid (NA). The
CC microspheres cause reduced inflammatory responses, and have increased
CC safety and biodegradability. The microspheres are useful for transfecting
CC a cell of a subject and for treating a subject having myocardial
CC ischemia, by increasing cardiac angiogenesis. They are also useful for
CC treating haemophilia. The present sequence represents the plasmid
CC pCDNA3.1/GS, into which is inserted a polynucleotide sequence which is
CC crosslinked to hyaluronic acid. The polynucleotide sequence encodes a
CC vascular endothelial growth factor (VEGF)
XX
SQ Sequence 4597 BP; 1062 A; 1214 C; 1206 G; 1115 T; 0 U; 0 Other;
Query Match 100.0%; Score 100; DB 4; Length 4597;
Best Local Similarity 100.0%; Pred. No. 5.5e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGGAGATCTCCGATCCCTATGCTCGACTCTCAGTACAAATCTGCTCTGATG 60
DB 1 GACGGATCGGGAGATCTCCGATCCCTATGCTCGACTCTCAGTACAAATCTGCTCTGATG 60

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100
DB 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100

RESULT 13
AAD39652

```

ID AAD39652 standard; DNA; 4639 BP.
XX
AC AAD39652;
XX
DT 22-OCT-2002 (first entry)
XX
DE Human small nuclear RNA (snRNA) DNA.
XX
KW Human; recombinant vector; insertion cassette; small nuclear RNA; snRNA;
KW transgenic animal; ds.
XX
OS Homo sapiens.
XX
PN US2002058287-A1.
XX
PD 16-MAY-2002.
XX
PF 12-MAR-2001; 2001US-0080481.
XX
PR 10-MAR-2000; 2000US-0188304P.
XX
PA (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX
PI Graaf DD, Lander ES;
XX
DR WPI; 2002-499510/53.
XX
PT New recombinant vector containing sequence for small nuclear RNA, useful
PT e.g. for identifying variant snRNA that suppresses expression of
PT transcription products.
XX
PS Disclosure; Fig 1; 18pp; English.
XX
CC The invention relates to a recombinant vector which comprises DNA,
CC consisting of an insertion cassette contained between at least two
CC insertion sites, that encodes a small nuclear (sn) RNA. The invention is
CC used to identify snRNA modifications that inhibit expression of
CC transcription products (and the identified snRNA are used to suppress
CC expression) for delivering antisense sequences to the nucleus and to
CC create transgenic animals. The present DNA sequence is human snRNA, U1
XX
SQ Sequence 4639 BP; 1067 A; 1198 C; 1243 G; 1131 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 6; Length 4639;
Best Local Similarity 100.0%; Pred. No. 5.5e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGATCGGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
DB 1 GACGATCGGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60

QY 61 CGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
DB 61 CGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

RESULT 14
AAF83146
ID AAF83146 standard; DNA; 4840 BP.
XX
AC AAF83146;
XX
DT 09-JUL-2001 (first entry)
XX
DE Complete sequence of vector pIRES/BS.
XX
KW Blastcidin resistance; BS gene; gene therapy; tissue engineering;
KW cosmetic surgery; arthritis; joint replacement; skin graft; rhinoplasty;
KW pIRES/BS; ss.
XX
OS Synthetic.
XX
PN WO200119853-A2.

AAD39652 standard; DNA; 4639 BP.
XX
AC AAD39652;
XX
DT 22-OCT-2002 (first entry)
XX
DE Human small nuclear RNA (snRNA) DNA.
XX
KW Human; recombinant vector; insertion cassette; small nuclear RNA; snRNA;
KW transgenic animal; ds.
XX
OS Homo sapiens.
XX
PN US2002058287-A1.
XX
PD 16-MAY-2002.
XX
PF 12-MAR-2001; 2001US-0080481.
XX
PR 10-MAR-2000; 2000US-0188304P.
XX
PA (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX
PI Graaf DD, Lander ES;
XX
DR WPI; 2002-499510/53.
XX
PT New recombinant vector containing sequence for small nuclear RNA, useful
PT e.g. for identifying variant snRNA that suppresses expression of
PT transcription products.
XX
PS Disclosure; Fig 1; 18pp; English.
XX
CC The invention relates to a recombinant vector which comprises DNA,
CC consisting of an insertion cassette contained between at least two
CC insertion sites, that encodes a small nuclear (sn) RNA. The invention is
CC used to identify snRNA modifications that inhibit expression of
CC transcription products (and the identified snRNA are used to suppress
CC expression) for delivering antisense sequences to the nucleus and to
CC create transgenic animals. The present DNA sequence is human snRNA, U1
XX
SQ Sequence 4639 BP; 1067 A; 1198 C; 1243 G; 1131 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 6; Length 4639;
Best Local Similarity 100.0%; Pred. No. 5.5e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGATCGGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
DB 1 GACGATCGGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60

QY 61 CGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
DB 61 CGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

RESULT 14
AAF83146
ID AAF83146 standard; DNA; 4840 BP.
XX
AC AAF83146;
XX
DT 09-JUL-2001 (first entry)
XX
DE Complete sequence of vector pIRES/BS.
XX
KW Blastcidin resistance; BS gene; gene therapy; tissue engineering;
KW cosmetic surgery; arthritis; joint replacement; skin graft; rhinoplasty;
KW pIRES/BS; ss.
XX
OS Synthetic.
XX
PN WO200119853-A2.
```

```

XX
PD 22-MAR-2001.
XX
XX
PF 11-SEP-2000; 2000WO-GB003462.
XX
XX
PR 11-SEP-1999; 99GB-00021418.
XX
XX
PA (UYSH-) UNIV SHEFFIELD.
XX
XX
PI Hollander AP, Barker MD, Kafienah W;
XX
XX
DR WPI; 2001-290354/30.
XX
XX
PT Novel nucleic acid molecule useful for therapeutic and cosmetic tissue
PT engineering, comprising at least a functional part of blastcidin
PT resistance gene linked through a recognition sequence, to a selected
PT gene.
XX
XX
PS Claim 11; Fig C; 44pp; English.
XX
CC The invention provides a nucleic acid molecule (I) comprising at least
CC the functional part of blastcidin resistance (BS) gene, or its homolog,
CC linked through a recognition sequence to at least one selected gene. (I)
CC is useful in treatment comprising: (1) providing cells/tissues transfected
CC with (I); (2) surgical administration of the cells/tissues to the patient
CC to be treated; and optionally (3) monitoring the status of the cells/
CC tissues by the patient. Therapeutic compositions comprising cells/tissues
CC transformed with (I) is useful in identifying the role of genes in
CC healthy and diseased tissue, in tissue engineering and in cosmetic
CC surgery. Tissue engineering can be used to treat arthritis, joint
CC replacement, skin grafts for burn victims, and replacement coronary
CC arteries. Cosmetic tissue surgery includes rhinoplasty. The present
CC sequence represents the nucleotide sequence of the vector pIRES/BS
CC containing the BS gene
XX
SQ Sequence 4840 BP; 1154 A; 1227 C; 1236 G; 1223 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 4; Length 4840;
Best Local Similarity 100.0%; Pred. No. 5.6e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGATCGGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
DB 1 GACGATCGGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60

QY 61 CGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
DB 61 CGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

RESULT 15
ADB33528
ID ADB33528 standard; DNA; 5015 BP.
XX
XX
AC ADB33528;
XX
XX
DT 04-DEC-2003 (first entry)
XX
DE Expression vector nucleotide sequence SEQ ID NO:27.
XX
KW fusion protein; amyloid precursor protein; APP; transcription factor;
KW neurotropic; neuroprotective; APP inhibitor;
KW amyloid precursor protein inhibitor; Alzheimer's disease; beta-secretase;
KW gamma-secretase; human; gene; ds.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO2003072041-A2.
XX
PD 04-SEP-2003.
XX
XX
PF 23-FEB-2003; 2003WO-US005458.
```

XX 27-FEB-2002; 2002US-0360274P.
XX (MERI) MERCK & CO INC.
XX
XX Espeseth AS, Ferrer M, Flores OA, Hazuda DJ, Inglesse J;
XX PI Miller MD, Register B, Shi X, Simon AU, Zuck PD;
XX WPI; 2003-689968/65.
XX
XX DNA encoding a fusion protein of amyloid precursor protein, useful in
XX screening for anti-Alzheimer agents, comprises a fused transcription
XX factor.
XX
XX Disclosure; Fig 32B-E; 193pp; English.
XX
XX The present invention describes a DNA molecule (I) that encodes a fusion
XX protein (FP) comprising: (i) an amino acid sequence of amyloid precursor
XX protein (APP), either the wild type, Swedish or NFEV versions; and (ii) a
XX transcription factor (TF), fused in frame to the C-terminus of (i). Also
XX described: (1) an expression vector containing (i); (2) a eukaryotic cell
XX containing (i); and (3) methods for identifying a compound (A) that
XX inhibits processing of APP, using the cells of (2). (i) has neurotropic and
XX neuroprotective activities. (i) can be used to produce eukaryotic cells
XX that express FP and are useful in screening for agents that inhibit
XX processing of APP. The agents are potentially useful for the treatment or
XX prevention of Alzheimer's disease. Cells that express FP can screen for
XX inhibitors of: (a) beta- and gamma-secretases; and (b)
XX cytoplasmic/extracellular APP signaling in a single assay. Cell-based
XX assays may be free of interference from alpha-secretase activity and are
XX homogeneous (no chromatography, immunoprecipitation or washing required)
XX so well suited to high-throughput screening. The present sequence
XX represents a plasmid nucleotide sequence from the present invention.
XX
SQ Sequence 5015 BP; 1167 A; 1297 C; 1279 G; 1272 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 10; Length 5015;
Best Local Similarity 100.0%; Pred. No. 5.6e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GACGGATCGGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTCCTCTGATG 60
Db 1 GACGGATCGGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTCCTCTGATG 60

Oy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100

Search completed: July 14, 2005, 07:01:37
Job time : 141.038 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 05:15:57 ; Search time 952.146 Seconds
(without alignments)
3997.736 Million cell updates/sec

Title: US-09-482-682-43_COPY_1_100

Perfect score: 100

Sequence: 1 gacggatcgaggagatcccc.....ctgtccctgctgtgtgtt 100

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsa1:*
9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	60	60.0	602	B67169	CpG0047A Cp
2	55.6	55.6	694	BZ052929	BZ052929 jnr13g03.
3	55.6	55.6	696	BZ050328	BZ050328 jnr42c12.
4	55.6	55.6	717	BZ054067	BZ054067 jnr38b09.
5	53.6	53.6	348	AW409112	AW409112 sal10h5 S
6	53.4	53.4	343	AL715724	AL715724 AL715724
7	53.4	53.4	345	AL714571	AL714571 AL714571
8	53.4	53.4	761	CK119397	CK119397 212c09.p1
9	53.4	53.4	766	CK120360	CK120360 207j04.p1
10	53.4	53.4	788	CK117844	CK117844 209p08.p1
11	53.4	53.4	898	CL141237	CL141237 ISB1-118J
12	53.4	53.4	899	CL140877	CL140877 ISB1-118B
13	53.4	53.4	1009	CL123953	CL123953 ISB1-84J1
14	53.2	53.2	814	AQ914559	AQ914559 nbe50049M
15	53	53.0	675	BZ051815	BZ051815 jnr57d03.
16	53	53.0	679	BZ052857	BZ052857 jnr13g03.
17	53	53.0	700	BZ050646	BZ050646 jnr66f08.
18	53	53.0	701	BZ052015	BZ052015 jnr56b03.
19	53	53.0	708	BZ054793	BZ054793 jnr33g03.
20	53	53.0	709	BZ053587	BZ053587 jnr98d01.
21	53	53.0	712	BZ054005	BZ054005 jnr38b09.
22	52.8	52.8	451	AQ863966	AQ863966 nbe50022E
23	52.6	52.6	399	AQ075099	AQ075099 CIT-HSP-2
24	52.4	52.4	700	BZ049113	BZ049113 jnr21d02.

25 52.4 52.4 708 8 BZ050047
26 51.6 51.6 328 9 CC819886
27 51.6 51.6 351 9 CC818492
28 51.6 51.6 358 9 CC817661
29 51.6 51.6 364 9 CC817805
30 51.6 51.6 364 9 CC818511
31 51.6 51.6 364 9 CC818574
32 51.6 51.6 364 9 CC819049
33 51.6 51.6 369 9 CC817069
34 51.6 51.6 374 9 CC817074
35 51.6 51.6 374 9 CC820036
36 51.6 51.6 395 9 CC817652
37 51.6 51.6 403 9 CC817682
38 51.6 51.6 403 9 CC817837
39 51.6 51.6 414 9 CC819240
40 51.6 51.6 419 9 CC818384
41 51.6 51.6 420 9 CC817834
42 51.6 51.6 426 9 CC817720
43 51.6 51.6 437 9 CC819820
44 51.6 51.6 441 9 CC818421
45 51.6 51.6 443 9 CC817769

ALIGNMENTS

RESULT 1
LOCUS B67169
DEFINITION CpG0047A CpTOWAGDNA2 Cryptosporidium parvum genomic, GSS 12-MAY-2000
sequence.
ACCESSION B67169
VERSION B67169.1 GI:2642750
KEYWORDS GSS.
SOURCE Cryptosporidium parvum
ORGANISM Cryptosporidium parvum
REFERENCE 1 (bases 1 to 602)
AUTHORS Strong, W.B. and Nelson, R.G.
TITLE Preliminary profile of the Cryptosporidium parvum genome: an expressed sequence tag and genome survey sequence analysis
JOURNAL Mol. Biochem. Parasitol. 107 (1), 1-32 (2000)
MEDLINE 20183851
PUBMED 10717299
COMMENT Contact: Nelson, R. G.
Depts. of Medicine & Pharmaceutical Chemistry
San Francisco General Hospital-University of California, San Francisco
Box 0811, San Francisco, CA 94143-0811, USA
Tel.: 415 206 8846
Fax: 415 206 3353
Email: malaria@itsa.ucsf.edu
Submitted sequence has been edited to remove vector sequences 5' to the insert, to correct miscalled bases and assign uncalled (N) bases throughout the sequence, and to terminate when base-calling became ambiguous.
Seq primer: T7
Class: Shotgun
High quality sequence stop: 602.
Location/Qualifiers
1. .602
/organism="Cryptosporidium parvum"
/mol_type="Genomic DNA"
/db_xref="taxon:5807"
/lab_host="E. coli XL2 Blue MRF"
/clone_lib="CpTOWAGDNA2"
/note="Vector: pCR-Script Amp SK+; Site_1: SrfI; C. parvum (IOWA isolate) genomic DNA was hydrodynamically sheared to produce fragments having a tight size distribution between 2-4 kb by Dr. Yvonne Thorntonsen of the Stanford DNA Sequencing and Technology Center

(http://sequence-www.stanford.edu/group/techdev/shear.htm)
 . The randomly sheared gDNA was chromatographed on Sephacryl S-400 to remove any small fragments and DNA eluting in the void volume was blunt-ended with T4 DNA Polymerase, treated with alkaline phosphatase to prevent the ligation of multiple fragments, ligated to Srf I-digested PCR-Script Amp (SK+) vector and transformed into E. coli strain XL10 Gold. Recombinant clones from the first plating of the library were selected for sequence analysis using T3 and T7 primers."

ORIGIN

Query Match 60.0%; Score 60; DB 8; Length 602;
 Best Local Similarity 100.0%; Pred. No. 2.4e-10;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 41 CAGTACAATCTGCTCTGATGCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTTT 100
 |||||
 Db 1 CAGTACAATCTGCTCTGATGCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTTT 60

RESULT 2

BZ052929/c
 LOCUS jnr13g03.g1 B.oleracea001 Brassica oleracea genomic, genomic survey
 DEFINITION
 ACCESSION BZ052929
 VERSION BZ052929.1 GI:23654922
 KEYWORDS GSS.
 SOURCE Brassica oleracea
 ORGANISM Brassica oleracea

REFERENCE
 AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

1 (bases 1 to 694)
 Delehaunty, K., Fewell, G., Fulton, L., McCombie, W.R., Miner, T.,

Nash, W., Rabinowicz, P.D. and Wilson, R.K.
 Whole genome shotgun reads from Brassica oleracea

Unpublished (2002)
 Contact: Richard K. Wilson

Genome Sequencing Center
 Washington University School of Medicine

Email: submissions@watson.wustl.edu
 Plate: jnr13 row: g column: 03

Seq primer: -28RPOT reverse

Class: shotgun

High quality sequence start: 32

High quality sequence stop: 551.

FEATURES

source

Location/Qualifiers
 1..694
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /db_xref="taxon:3712"
 /clone_lib="B.oleracea001"
 /note="Vector: pOTw13; Whole genome shotgun library from flowering buds. DNA was purified from a crude nuclear prep using Brassica oleracea T01000DH3 buds provided by Thomas Osborn at the University of Wisconsin. Genomic DNA was provided by Pablo Rabinowicz (CSHL) and the shotgun library prepared at Washington University Genome Sequencing Center."

ORIGIN

Query Match 55.6%; Score 55.6; DB 8; Length 694;
 Best Local Similarity 77.9%; Pred. No. 9e-09;
 Matches 67; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 3 CGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATGCC 62
 |||||
 Db 324 CGGATCGATAGTCCCTCGACTAGTTATGTCGACTCTCAGTACAATCTGCTCTGATGCC 265
 |||||
 Qy 63 GCATAGTTAAGCCAGTATCTGCTCC 88
 |||||

Db 264 GCATAGTTAAGCCAGCCCGACACCC 239

RESULT 3

BZ050328
 LOCUS jnr42c12.b1 B.oleracea001 Brassica oleracea genomic, genomic survey
 DEFINITION
 ACCESSION BZ050328
 VERSION BZ050328.1 GI:23649718
 KEYWORDS GSS.
 SOURCE Brassica oleracea
 ORGANISM Brassica oleracea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

1 (bases 1 to 696)

Delehaunty, K., Fewell, G., Fulton, L., McCombie, W.R., Miner, T.,

Nash, W., Rabinowicz, P.D. and Wilson, R.K.
 Whole genome shotgun reads from Brassica oleracea

Unpublished (2002)

Contact: Richard K. Wilson
 Genome Sequencing Center

Washington University School of Medicine
 Email: submissions@watson.wustl.edu

Plate: jnr42 row: c column: 12
 Seq primer: -21UPPOT forward

Class: shotgun

High quality sequence start: 35

High quality sequence stop: 180.

FEATURES

source

Location/Qualifiers
 1..696
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /db_xref="taxon:3712"
 /clone_lib="B.oleracea001"
 /note="Vector: pOTw13; Whole genome shotgun library from flowering buds. DNA was purified from a crude nuclear prep using Brassica oleracea T01000DH3 buds provided by Thomas Osborn at the University of Wisconsin. Genomic DNA was provided by Pablo Rabinowicz (CSHL) and the shotgun library prepared at Washington University Genome Sequencing Center."

ORIGIN

Query Match 55.6%; Score 55.6; DB 8; Length 696;
 Best Local Similarity 77.9%; Pred. No. 9e-09;
 Matches 67; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 3 CGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATGCC 62
 |||||
 Db 45 CGGATCGATAGTCCCTCGACTAGTTATGTCGACTCTCAGTACAATCTGCTCTGATGCC 104
 |||||
 Qy 63 GCATAGTTAAGCCAGTATCTGCTCC 88
 |||||
 Db 105 GCATAGTTAAGCCAGCCCGACACCC 130

RESULT 4

BZ054067/c
 LOCUS jnr38b09.g1 B.oleracea001 Brassica oleracea genomic, genomic survey
 DEFINITION
 ACCESSION BZ054067
 VERSION BZ054067.1 GI:23657216
 KEYWORDS GSS.
 SOURCE Brassica oleracea
 ORGANISM Brassica oleracea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

1 (bases 1 to 717)

Delehaunty, K., Fewell, G., Fulton, L., McCombie, W.R., Miner, T.,

Site 2: NotI; About 1 week after bolting, cDNA synthesis using SuperscriptTM-system (Invitrogen) with an oligo(dT)-primer containing NotI restriction site and a SalI adapter. The main library (plate numbers begin with 1) of 38,000 clones was rearrayed into the sublibrary (plate numbers begin with 201) containing 5,000 putative expression clones. Average insert size is 1 kb. Note: The rearrayed sublibrary (plate numbers begin with 201) was sequenced. Library generation and sequencing was granted in context of GABI-LAPP; data are also accessible at <https://gabi.rzpd.de>

ORIGIN

Query Match 53.4%; Score 53.4; DB 7; Length 766;
Best Local Similarity 84.5%; Pred. No. 5.6e-08;
Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 17 TCCGATCCCTATGCTGCTCAGTACATCTGCTGTGATCGCGCATAGTTAAGCCA 76
Db 679 TTCACCGCATATGGTGCACTCTCAGTACATCTGCTGTGATCGCGCATAGTTAAGCCA 620

Qy 77 GTATCTGCTCC 87
Db 619 GTATACACTCC 609

RESULT 10

CK117844/c

LOCUS CK117844 788 bp mRNA linear EST 01-JUN-2004
DEFINITION 209p08.p1 Atm1 Arabidopsis thaliana cDNA clone MPMP2011P08209
5-PRIME, mRNA sequence.

ACCESSION CK117844

VERSION CK117844.1

KEYWORDS EST. GI:47828160

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 788)

REFERENCE Feilner, T., Immink, R.G.H., Cahill, D.J. and Kersten, B.

Generation of a cDNA expression library from Arabidopsis

inflorescence meristem

Unpublished (2003)

CONTACT Birgit Kersten

Plant Protein Chip Group, Department Lehrach

Max-Planck-Institute for Molecular Genetics

Innestr. 73, D-14195 Berlin, Germany

Tel: +49(0)30/84131648

Fax: +49(0)30/84131128

Email: Kersten@molgen.mpg.de

Insert Length: 788 Std Error: 0.00

Plate: 209 row: P column: 8

Seq primer: pQE65.

Location/Qualifiers

1..788

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/ecotype="Columbia"

/db_xref="GABI:953578"

/db_xref="taxon:3702"

/clone="MPMP2011P08209"

/tissue_type="inflorescence meristem"

/dev_stage="about one week after bolting"

/lab_host="E. coli SCS-1/pSE111"

/clone_lib="AtM1"

/note="Vector: pQS-3ONAST-attB (AY386205); Site 1: SalI;

Site 2: NotI; About 1 week after bolting, cDNA synthesis

using SuperscriptTM-system (Invitrogen) with an

oligo(dT)-primer containing NotI restriction site and a

SalI adapter. The main library (plate numbers begin with

1) of 38,000 clones was rearrayed into the sublibrary

(plate numbers begin with 201) containing 5,000 putative

expression clones. Average insert size is 1 kb. Note: The rearrayed sublibrary (plate numbers begin with 201) was sequenced. Library generation and sequencing was granted in context of GABI-LAPP; data are also accessible at <https://gabi.rzpd.de>

ORIGIN

Query Match 53.4%; Score 53.4; DB 7; Length 788;
Best Local Similarity 84.5%; Pred. No. 5.7e-08;
Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 17 TCCGATCCCTATGCTGCTCAGTACATCTGCTGTGATCGCGCATAGTTAAGCCA 76

Db 514 TTCACCGCATATGGTGCACTCTCAGTACATCTGCTGTGATCGCGCATAGTTAAGCCA 455

Qy 77 GTATCTGCTCC 87

Db 454 GTATACACTCC 444

RESULT 11

CL141237/c

LOCUS CL141237 898 bp DNA linear GSS 05-JAN-2004

DEFINITION ISB1-118J17_T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-118J17,
genomic survey sequence.

ACCESSION CL141237

VERSION CL141237.1

KEYWORDS GSS. GI:40634872

SOURCE Xenopus tropicalis (western clawed frog)

ORGANISM Xenopus tropicalis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;

Xenopodinae; Xenopus; Silurana.

1 (bases 1 to 898)

REFERENCE Kremitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,

Mardis, E. and Wilson, R.

A physical map of the xenopus tropicalis genome

Unpublished (2003)

CONTACT Richard K Wilson

Genome Sequencing Center

Washington University School of Medicine

Email: submissions@wustl.edu

Insert Length: 75000 Std Error: 0.00

Seq primer: T7 TAATACGACTCACTATAGG

Class: BAC ends

High quality sequence start: 4

High quality sequence stop: 742.

Location/Qualifiers

1..898

/organism="Xenopus tropicalis"

/mol_type="genomic DNA"

/db_xref="taxon:8364"

/clone="ISB1-118J17"

/clone_lib="ISB1"

/note="vector: pBelOBAC11; ISB-1 Xenopus tropicalis BAC

Library Segment 1"

ORIGIN

Query Match 53.4%; Score 53.4; DB 9; Length 898;
Best Local Similarity 84.5%; Pred. No. 5.8e-08;
Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 17 TCCGATCCCTATGCTGCTCAGTACATCTGCTGTGATCGCGCATAGTTAAGCCA 76

Db 195 TTCACCGCATATGGTGCACTCTCAGTACATCTGCTGTGATCGCGCATAGTTAAGCCA 136

Qy 77 GTATCTGCTCC 87

Db 135 GTATACACTCC 125

RESULT 12

CL140877/c

```

LOCUS       CL140877               899 bp    DNA    linear    GSS 05-JAN-2004
DEFINITION   ISB1-118B12_T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-118B12,
              genomic survey sequence.
ACCESSION    CL140877
VERSION      CL140877.1    GI:40634512
KEYWORDS     GSS.
SOURCE       Xenopus tropicalis (western clawed frog)
ORGANISM     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
              Xenopodinae; Xenopus; Silurana.
REFERENCE    1 (bases 1 to 899)
AUTHORS      Krenitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
              Mardis,E. and Wilson,R.
TITLE        A physical map of the xenopus tropicalis genome
JOURNAL      Unpublished (2003)
COMMENT      Contact: Richard K Wilson
              Genome Sequencing Center
              Washington University School of Medicine
              Email: submissions@watson.wustl.edu
              Insert Length: 75000 Std Error: 0.00
              Seq primer: T7 TAATACGACTCTACTATAGGG
              Class: BAC ends
              High quality sequence start: 4
              High quality sequence stop: 681.
              Location/Qualifiers
                1..899
                /organism="Xenopus tropicalis"
                /mol_type="genomic DNA"
                /db_xref="taxon:8364"
                /clone="ISB1-118B12"
                /clone_lib="ISB1"
                /note="Vector: pBelOBAC11; ISB-1 Xenopus tropicalis BAC
                  Library Segment 1"

FEATURES             source
  source              1..899
  Query Match         53.4%; Score 53.4; DB 9; Length 899;
  Best Local Similarity 84.5%; Pred. No. 5.8e-08;
  Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 17 TCCGATCCCTATGTCGACTCTCAGTACAACTGCTGATGCCGATAGTTAAGCCA 76
Db 195 TTCACACCGCATATGGTGCATCTCAGTACAACTGCTGATGCCGATAGTTAAGCCA 136

Qy 77 GTATCTGCTCC 87
Db 135 GTATACACTCC 125

RESULT 13
LOCUS       CL123953/c
DEFINITION   ISB1-84J15_T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-84J15,
              genomic survey sequence.
ACCESSION    CL123953
VERSION      CL123953.1    GI:40617588
KEYWORDS     GSS.
SOURCE       Xenopus tropicalis (western clawed frog)
ORGANISM     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
              Xenopodinae; Xenopus; Silurana.
REFERENCE    1 (bases 1 to 1009)
AUTHORS      Krenitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
              Mardis,E. and Wilson,R.
TITLE        A physical map of the xenopus tropicalis genome
JOURNAL      Unpublished (2003)
COMMENT      Contact: Richard K Wilson
              Genome Sequencing Center
              Washington University School of Medicine
              Email: submissions@watson.wustl.edu
              Insert Length: 75000 Std Error: 0.00
              Seq primer: T7 TAATACGACTCTACTATAGGG

```

```

Class: BAC ends
High quality sequence start: 167
High quality sequence stop: 324.
Location/Qualifiers
  1..1009
  /organism="Xenopus tropicalis"
  /mol_type="genomic DNA"
  /db_xref="taxon:8364"
  /clone="ISB1-84J15"
  /clone_lib="ISB1"
  /note="Vector: pBelOBAC11; ISB-1 Xenopus tropicalis BAC
    Library Segment 1"

ORIGIN
Query Match         53.4%; Score 53.4; DB 9; Length 1009;
Best Local Similarity 84.5%; Pred. No. 5.9e-08;
Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 17 TCCGATCCCTATGTCGACTCTCAGTACAACTGCTGATGCCGATAGTTAAGCCA 76
Db 252 TTCACACCGCATATGGTGCATCTCAGTACAACTGCTGATGCCGATAGTTAAGCCA 193

Qy 77 GTATCTGCTCC 87
Db 192 GTATACACTCC 182

RESULT 14
LOCUS       AQ914559               814 bp    DNA    linear    GSS 02-DEC-1999
DEFINITION   nbeb0049M21r CUGI Rice BAC Library (ECORI) Oryza sativa (japonica
              cultivar-group) genomic clone nbeb0049M21r, genomic survey
              sequence.
ACCESSION    AQ914559
VERSION      AQ914559.1    GI:6511075
KEYWORDS     GSS.
SOURCE       Oryza sativa (japonica cultivar-group)
              Oryza sativa (japonica cultivar-group)
              Rukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
              Ehrhartoideae; Oryzaceae; Oryza.
              1 (bases 1 to 814)
              Wing,R.A. and Dean,R.A.
              A BAC End Sequencing Framework to Sequence the Rice Genome
              Unpublished (1998)
              Contact: Wing RA
              Clemson University Genomics Institute
              Clemson University
              100 Jordan Hall, Clemson, SC 29634, USA
              Tel: 864 656 7288
              Fax: 864 656 4293
              Email: rwing@clemson.edu
              Seq primer: GGAACAGCTATGACCATG
              Class: BAC ends
              High quality sequence start: 21
              High quality sequence stop: 361.
              Location/Qualifiers
                1..814
                /organism="Oryza sativa (japonica cultivar-group)"
                /mol_type="genomic DNA"
                /cultivar="Nipponbare"
                /db_xref="taxon:39947"
                /clone="nbeb0049M21r"
                /tissue_type="Leaf"
                /lab_host="E. coli DH10B"
                /clone_lib="CUGI Rice BAC Library (ECORI)"
                /note="vector: pBACindigo; Site:1: EcoRI; Site:2: EcoRI;
                  Rice is the most important food crop in the world. Half of
                  the world population, especially those inhabiting highly
                  populated areas of the humid tropics and subtropics, rely
                  on rice as their primary source of carbohydrate.
                  Monocotyledonous rice is a diploid plant (2n=24) with a
                  haploid genome equivalent of 431 Mbp (Arumuganathan and

```

Barle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from *Oryza sativa*, Nipponbare variety using EcoRI as the cloning enzyme. The library contains 55,296 clones with an average insert size of 121 kb providing approximately 15 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9%. Three high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening and can be requested from the Clemson University BAC/EST Resource Center (www.genome.clemson.edu)."

ORIGIN

Query Match 53.2%; Score 53.2; DB 8; Length 814;
 Best Local Similarity 78.0%; Pred. No. 6.7e-08;
 Matches 64; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 7 TCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATGCCGAT 66
 |||||
 Db 279 TGGGCGGATTTTCACACCGCATATGTCGACTCTCAGTACAAATCTGCTCTGATGCCGAT 338
 |||||

Qy 67 AGTTAAGCCAGTATCTGCTCCC 88
 |||||

Db 339 AGTTAAGCCAGCCCGCACCC 360
 |||||

RESULT 15

BZ051815
 LOCUS jnr57d03.b1 B.oleracea001 675 bp DNA linear GSS 09-OCT-2002
 DEFINITION sequence.
 BZ051815
 VERSION BZ051815.1 GI:23652690
 KEYWORDS GSS.
 SOURCE Brassica oleracea
 ORGANISM Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 675)
 Delehaunty,K., Fewell,G., Fulton,L., McCombie,W.R., Miner,T.,
 Nash,W., Rabinowicz,P.D. and Wilson,R.K.
 Whole genome shotgun reads from Brassica oleracea
 Unpublished (2002)
 Contact: Richard K. Wilson
 Genome Sequencing Center
 Washington University School of Medicine
 Email: submissions@watson.wustl.edu
 Plate: jnr57 row: d column: 03
 Seq primer: -21UPpOT forward
 Class: shotgun
 High quality sequence start: 29
 High quality sequence stop: 94.

FEATURES

source
 1..675
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /db_xref="taxon:3712"
 /clone_lib="B.oleracea001"
 /note="Vector: pOTw13; Whole genome shotgun library from
 flowering buds. DNA was purified from a crude nuclear
 prep using Brassica oleracea T01000DH3 buds provided by
 Thomas Osborn at the University of Wisconsin. Genomic
 DNA was provided by Pablo Rabinowicz (CSHL) and the
 shotgun library prepared at Washington University Genome
 Sequencing Center."

ORIGIN

Query Match 53.0%; Score 53; DB 8; Length 675;
 Best Local Similarity 75.6%; Pred. No. 7.6e-08;

Matches 65; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy 3 CGGATCGGGAGATCTCCCGATCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATGCC 62
 |||||
 Db 53 CGGNACGATAGGTCCCTGGACTAGTTATGTTGCACTCTCAGTACAAATCTGCTCTGATGCC 112
 |||||

Qy 63 GCATAGTTAAGCCAGTATCTGCTCCC 88
 |||||

Db 113 GCATAGTTAAGCCAGCCCGCACCC 138
 |||||

Search completed: July 14, 2005, 23:22:51
 Job time : 952.146 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:39:07 ; Search time 756.618 Seconds
(without alignments)
6468.225 Million cell updates/sec

Title: US-09-482-682-43_COPY_4141_4241

Perfect score: 101
Sequence: 1 acagcaagggggaggattgg.....ccagctggggctctaggggg 101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_hg.*
- 3: gb_in.*
- 4: gb_on.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	3986	12	PCDNA3ZEO
2	101	100.0	4597	6	AX060344 Sequence
3	101	100.0	5082	6	A91754 Sequence 10
4	101	100.0	5082	6	BD085110 Vertebrat
5	101	100.0	5432	6	BD234590 Screening
6	101	100.0	5432	6	BD026821 Sequence
7	101	100.0	5446	6	BD195386 Compositi
8	101	100.0	5446	6	AX319694 Sequence
9	101	100.0	5590	12	AB038602 Cloning v
10	101	100.0	5639	12	AX437643 Expressio
11	101	100.0	5651	6	AX211282 Sequence
12	101	100.0	5651	6	AX349366 Sequence
13	101	100.0	5731	6	AX202478 Sequence
14	101	100.0	5995	6	AX685746 Sequence
15	101	100.0	6084	12	CGA575208
16	101	100.0	6109	12	TRU90717
17	101	100.0	6148	6	BD181637
18	101	100.0	6148	6	AX342685 Sequence
19	101	100.0	6149	6	BD181638 Novel mel

20	101	100.0	6149	6	AX342686
21	101	100.0	6180	6	AX207724
22	101	100.0	6186	6	AX211281
23	101	100.0	6186	6	AX349365
24	101	100.0	6195	6	BD168975
25	101	100.0	6213	6	AX211283
26	101	100.0	6213	6	AX349369
27	101	100.0	6238	6	BD168966
28	101	100.0	6253	6	AR031374
29	101	100.0	6253	6	BD009742
30	101	100.0	6277	12	AV437644
31	101	100.0	6331	12	EVPCMVPA1
32	101	100.0	6333	12	EVPCMVPA3
33	101	100.0	6334	6	AX665478
34	101	100.0	6335	12	EVPCMVPA2
35	101	100.0	6338	6	BD134374
36	101	100.0	6338	6	AR428934
37	101	100.0	6340	6	AX207733
38	101	100.0	6365	6	AX513181
39	101	100.0	6394	12	AF416990
40	101	100.0	6404	6	BD267665
41	101	100.0	6411	6	AX207725
42	101	100.0	6411	6	AX207729
43	101	100.0	6420	6	BD267666
44	101	100.0	6436	6	AX207740
45	101	100.0	6439	6	AR240214

ALIGNMENTS

RESULT 1	PCDNA3ZEO	3986 bp	DNA	linear	SVN 16-AUG-1995
LOCUS	Cloning vector pcdna3zEO	DNA			
DEFINITION	X90639				
ACCESSION	X90639.1	GI:949972			
VERSION					
KEYWORDS	cloning vector; expression vector; multiple cloning site; Plasmid.				
SOURCE	synthetic construct				
ORGANISM	other sequences; artificial sequences.				
REFERENCE	1				
AUTHORS	Peters, H., Hundhausen, T., Kroenke, M. and Marget, M.				
TITLE	A new small sized high-level eukaryotic expression vector				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 3986)				
AUTHORS	Peters, H.				
TITLE	Direct Submission				
JOURNAL	Submitted (07-AUG-1995) H. Peters, Inst. f. Immunologie,				
COMMENT	Michaelistr. 5, D- 24105 Kiel, FRG				
FEATURES	Related sequences: M21295 and K03104.				
source	Location/Qualifiers				
	1..3986				
	/organism="synthetic construct"				
	/mol_type="other DNA"				
	/db_xref="taxon:32630"				
	/plasmid="pcDNA3ZEO"				
misc_feature	1..2125				
	/note="cloning vector (pcDNA3) (Invitrogen)"				
misc_feature	889..994				
	/note="multiple cloning site (MCS)"				
misc_feature	2126..2796				
	/note="cloning vector (PZeoSV) (Invitrogen)"				
misc_feature	2797..3986				
	/note="cloning vector (pcDNA3)"				

ORIGIN

Query Match 100.0%; Score 101; DB 12; Length 3986;
Best Local Similarity 100.0%; Pred. No. 2.5e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 ACAGCAAGGGGAGGATTGGGAGACATAGCAGCATCTGGGGATCGGGGCTCTA 60
|||||

```

Db      1204 ACAGCAAGGGGAGGATTGGGAAGCAATAGCAGCATGCTGGGATCGGTGGGCTCTA 1263
Qy      61 TGGCTTCTGAGCGGGAAGAACCAAGCTGGGCTCTAGGGG 101
        |||||
Db      1264 TGGCTTCTGAGCGGGAAGAACCAAGCTGGGCTCTAGGGG 1304

RESULT 2
AX060344
LOCUS      AX060344
DEFINITION Sequence 3 from Patent WO078358.
ACCESSION AX060344
VERSION    AX060344.1 GI:12405832
KEYWORDS
SOURCE     synthetic construct
ORGANISM   other sequences; artificial sequences.
REFERENCE 1
AUTHORS    Chen, W.
TITLE      Hyaluronic acid microspheres for sustained gene transfer
JOURNAL    Patent: WO 0078358-A 3 28-DEC-2000;
           The Collaborative Group, Ltd. (US)
FEATURES   Location/Qualifiers
            source
              1. 4597
                /organism="synthetic construct"
                /mol_type="unassigned DNA"
                /db_xref="taxon:32630"
                /note="pCDNA3.1/GS vector by Invitrogen Corporation"
ORIGIN
Query Match      100.0%; Score 101; DB 6; Length 4597;
Best Local Similarity 100.0%; Pred. No. 2.5e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ACAGCAAGGGGAGGATTGGGAAGCAATAGCAGCATGCTGGGATCGGTGGGCTCTA 60
        |||||
Db      1780 ACAGCAAGGGGAGGATTGGGAAGCAATAGCAGCATGCTGGGATCGGTGGGCTCTA 1839

Qy      61 TGGCTTCTGAGCGGGAAGAACCAAGCTGGGCTCTAGGGG 101
        |||||
Db      1840 TGGCTTCTGAGCGGGAAGAACCAAGCTGGGCTCTAGGGG 1880

RESULT 3
A91754
LOCUS      A91754
DEFINITION Sequence 10 from Patent WO9824810.
ACCESSION  A91754
VERSION     A91754.1 GI:6740671
KEYWORDS
SOURCE     unidentified
ORGANISM   unclassified.
REFERENCE 1 (bases 1 to 5082)
AUTHORS    Bogaert, T.A. and Deraeymaeker, M.
TITLE      VERTEBRATE HOMOLOGUES OF UNC-53 PROTEIN OF C. ELEGANS
JOURNAL    Patent: WO 9824810-A 10 11-JUN-1998;
           BOGAERT THIERRY ANDRE OLIVIER (BE); DERAEMYAEKER MARC (BE)
FEATURES   Location/Qualifiers
            source
              1. 5082
                /organism="unidentified"
                /mol_type="unassigned DNA"
                /db_xref="taxon:32644"
ORIGIN
Query Match      100.0%; Score 101; DB 6; Length 5082;
Best Local Similarity 100.0%; Pred. No. 2.5e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ACAGCAAGGGGAGGATTGGGAAGCAATAGCAGCATGCTGGGATCGGTGGGCTCTA 60
        |||||
Db      2874 ACAGCAAGGGGAGGATTGGGAAGCAATAGCAGCATGCTGGGATCGGTGGGCTCTA 2933

Db      2874 ACAGCAAGGGGAGGATTGGGAAGCAATAGCAGCATGCTGGGATCGGTGGGCTCTA 2933

```

```

Qy      61 TGGCTTCTGAGCGGGAAGAACCAAGCTGGGCTCTAGGGG 101
        |||||
Db      2934 TGGCTTCTGAGCGGGAAGAACCAAGCTGGGCTCTAGGGG 2974

RESULT 4
BD085110
LOCUS      BD085110
DEFINITION Vertebrate homologues of UNC-53 protein of C elegans.
ACCESSION  BD085110
VERSION     BD085110.1 GI:22630720
KEYWORDS    JP 2001522222-A/8.
SOURCE      unidentified
ORGANISM    unclassified.
REFERENCE 1 (bases 1 to 5082)
AUTHORS    Platteeuw, C.J., Arjol, C.M.B., Deraeymaeker, M., Verhasselt, P.,
           Pujol, N.J.R., Luc, Maertens, J.S., Luyten, W., Geerts, H.,
           Vandekerckhove, J.S., Geysen, J. and Bogaert, T.A.O.E.
TITLE      Vertebrate homologues of UNC-53 protein of C elegans
JOURNAL    Patent: JP 2001522222-A 8 13-NOV-2001;
           JANSSEN PHARMACEUTICA NV
FEATURES   Location/Qualifiers
            OS Unidentified
            PN JP 2001522222-A/8
            PD 13-NOV-2001
            PF 03-DEC-1997 JP 1998525231
            PR 04-DEC-1996 GB 9625283.8
            PI CHRIST JULES PLATTEEUW, CARLOS MANUEL BUESA ARJOL, MARC PI
              DERAEMYAEKER,
              PI PETER VERHASSELT, NATHALIE JEANNE RAYMONDE PUJOL, LUC PI
              JACQUES SIMON MAERTENS,
              PI WALTER LUYTEN, HUGO GEERTS, JOEL STEFAAN VANDEKERCKHOVE, JOHAN
              PI GEYSEN,
              PI THIERRY ANDRE OLIVIER EDDY BOGAERT
            PC C12N15/12, C12N5/10, C12N15/85, C07K14/435, C07K16/18, A61K38/17,
            PC A61K49/00,
            PC C12Q1/02, G01N33/53
            CC Strandedness: Double;
            CC Topology: Circular;
            CC Vertebrate homologues of UNC-53 protein of C elegans FH Key
            Location/Qualifiers
            FT source
              1. 5082
                /organism="Unidentified".
            Location/Qualifiers
            source
              1. 5082
                /organism="unidentified"
                /mol_type="genomic DNA"
                /db_xref="taxon:32644"
ORIGIN
Query Match      100.0%; Score 101; DB 6; Length 5082;
Best Local Similarity 100.0%; Pred. No. 2.5e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ACAGCAAGGGGAGGATTGGGAAGCAATAGCAGCATGCTGGGATCGGTGGGCTCTA 60
        |||||
Db      2874 ACAGCAAGGGGAGGATTGGGAAGCAATAGCAGCATGCTGGGATCGGTGGGCTCTA 2933

Qy      61 TGGCTTCTGAGCGGGAAGAACCAAGCTGGGCTCTAGGGG 101
        |||||
Db      2934 TGGCTTCTGAGCGGGAAGAACCAAGCTGGGCTCTAGGGG 2974

RESULT 5
BD234590
LOCUS      BD234590
DEFINITION Screening assay of Abeta-peptide.
ACCESSION  BD234590
VERSION     BD234590.1 GI:33044360
KEYWORDS    JP 2002531141-A/2.
SOURCE      synthetic construct
ORGANISM    other sequences; artificial sequences.

```


REFERENCE 1 (bases 1 to 5432)
AUTHORS Peraus,G.
TITLE Screening assay of Abeta-peptide
JOURNAL Patent: JP 2002531141-A 2 24-SEP-2002;
COMMENT AVENTIS PHARMA DEUTSCHLAND GMBH
OS Artificial Sequence
PN JP 2002531141-A/2
PD 24-SEP-2002
PF 27-NOV-1999 JP 2000586944
PR 07-DEC-1998 DE 198 56 261.6
PI GISELA PERAUS
PC C12N15/09,A01K67/033,A61K45/00,A61P25/28,C12N1/15,C12N1/19, PC
C12N1/21,
PC C12N5/10,C12Q1/37,C12Q1/68,C12N15/00,C12N5/00 CC Description
of Artificial Sequence: Mutagen
FH Key Location/Qualifiers
FT source 1..5432
FT /organism='Artificial Sequence'.
FEATURES
source
1..5432
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 5432;
Best Local Similarity 100.0%; Pred. No. 2.4e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ACAGCAAGGGGAGGATTGGGAAGACATAGCAGCATCTGGGGATCGGTGGGCTCTA 60
Db 1190 ACAGCAAGGGGAGGATTGGGAAGACATAGCAGCATCTGGGGATCGGTGGGCTCTA 1249
Qy 61 TGGCTTCTGAGCGGAAGAACACCTGCGGCTCTAGGGGG 101
Db 1250 TGGCTTCTGAGCGGAAGAACACCTGCGGCTCTAGGGGG 1290
RESULT 6
LOCUS AX026821 5432 bp DNA linear PAT 16-SEP-2000
DEFINITION Sequence 9 from Patent DE19856261.
ACCESSION AX026821
VERSION AX026821.1 GI:10187947
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE 1
AUTHORS Peraus,G.
JOURNAL Patent: DE 19856261-C 9 30-MAR-2000;
HOBCHST MARION ROUSSEL DE GMBH (DE)
FEATURES
source
1..5432
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Mutagen"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 5432;
Best Local Similarity 100.0%; Pred. No. 2.4e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ACAGCAAGGGGAGGATTGGGAAGACATAGCAGCATCTGGGGATCGGTGGGCTCTA 60
Db 1190 ACAGCAAGGGGAGGATTGGGAAGACATAGCAGCATCTGGGGATCGGTGGGCTCTA 1249
Qy 61 TGGCTTCTGAGCGGAAGAACACCTGCGGCTCTAGGGGG 101
Db 1250 TGGCTTCTGAGCGGAAGAACACCTGCGGCTCTAGGGGG 1290

RESULT 7
LOCUS BD195386 5446 bp DNA linear PAT 17-JUL-2003
DEFINITION Composition and methods for administering Pneumococcal DNA.
ACCESSION BD195386
VERSION BD195386.1 GI:33005156
KEYWORDS JP 2002514061-A/3.
SOURCE unidentified
ORGANISM unidentified
unclassified.
REFERENCE 1 (bases 1 to 5446)
AUTHORS Briles,D.E., Mcdaniel,L.S. and Curiel,D.T.
TITLE Composition and methods for administering Pneumococcal DNA
JOURNAL Patent: JP 2002514061-A 3 14-MAY-2002;
UNIVERSITY OF ALABAMA AT BIRMINGHAM
COMMENT OS Unidentified
PN JP 2002514061-A/3
PD 14-MAY-2002
PF 04-DEC-1997 JP 1998535895
PR 04-DEC-1996 US 08/759505
PI DAVID E BRILES, LARRY S MCDANIEL, DAVID T CURIEL PC
C12P21/06,C12N15/00,C07H21/02,C07H21/04
CC Strandedness: Single;
CC Topology: Linear;
CC Composition and methods for administering Pneumococcal DNA FH
Key Location/Qualifiers
FT source 1..5446
FT /organism='Unidentified'.
FEATURES
source
1..5446
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 5446;
Best Local Similarity 100.0%; Pred. No. 2.4e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ACACCAAGGGGAGGATTGGGAAGACATAGCAGCATCTGGGGATCGGTGGGCTCTA 60
Db 1204 ACACCAAGGGGAGGATTGGGAAGACATAGCAGCATCTGGGGATCGGTGGGCTCTA 1263
Qy 61 TGGCTTCTGAGCGGAAGAACACCTGCGGCTCTAGGGGG 101
Db 1264 TGGCTTCTGAGCGGAAGAACACCTGCGGCTCTAGGGGG 1304
RESULT 8
LOCUS AX319694 5446 bp DNA linear PAT 14-DEC-2001
DEFINITION Sequence 5 from Patent WO0181614.
ACCESSION AX319694
VERSION AX319694.1 GI:17901350
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE 1
AUTHORS Leng,J.
TITLE Cell proliferation assay
JOURNAL Patent: WO 0181614-A 5 01-NOV-2001;
Chemicon International (US)
FEATURES
source
1..5446
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="pcDNA3 vector sequence"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 5446;
Best Local Similarity 100.0%; Pred. No. 2.4e-18;

```
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGCATGCTGGGATGCGGTGGGCTCTA 60
|
|
|
Db 1204 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGCATGCTGGGATGCGGTGGGCTCTA 1263
|
|
|
Qy 61 TGGCTTCTGAGCGGGAAGACCAAGCTGGGCTCTAGGGG 101
|
|
|
Db 1264 TGGCTTCTGAGCGGGAAGACCAAGCTGGGCTCTAGGGG 1304
|
|
|

RESULT 9
AB038602/c
LOCUS AB038602 5590 bp DNA circular SYN 24-JAN-2004
DEFINITION Cloning vector pLCPVRGNR104 DNA, complete sequence.
ACCESSION AB038602
VERSION AB038602.1 GI:13094141
KEYWORDS Cloning vector pLCPVRGNR104
SOURCE Cloning vector pLCPVRGNR104
ORGANISM other sequences; artificial sequences; vectors.
REFERENCE 1 Hashinaka,K.
AUTHORS Hashinaka,K.
TITLE Synthetic Autonomous Vectors Based on Palindromic Sequences of Parvovirus B19
JOURNAL Published Only in Database (2001)
REFERENCE 2 (bases 1 to 5590)
AUTHORS Hashinaka,K.
TITLE Direct Submision
JOURNAL Submitted (21-FEB-2000) Kazuya Hashinaka, Miyazaki Medical College, Department of Biochemistry, 5200 Kihara, Kiyotake, Miyazaki 889-1692, Japan (E-mail:hashinaka@post1.miyazaki-med.ac.jp, Tel:81-985-85-0985, Fax:81-985-85-2401)

FEATURES
source
1. 5590
/organism="Cloning vector pLCPVRGNR104"
/mol_type="other DNA"
/db_xref="taxon:117920"
/focus
12. 485
/organism="B19 virus"
/mol_type="other DNA"
/db_xref="taxon:10798"
/notes="synonym:Parvovirus B19"
3369. 3846
/organism="B19 virus"
/mol_type="other DNA"
/db_xref="taxon:10798"
/notes="synonym:Parvovirus B19"

repeat_region
gene
15. 397
complement(1115..2635)
/gene="rsgFP-neor"
CDS
complement(1115..2635)
/gene="rsgFP-neor"
/codon_start=1
/transl_table=11
/product="red-shift green fluorescent protein-fused neomycin phosphotransferase"
/protein_id="BAB32740.1"
/db_xref="GI:13094142"
/translacion="MASKGELFTGVVPILVELDGDVNGHKFVSQGEQDATYVKGLT LKPICTGKLPVMPITLTTCYQVCFSPRYPDMKHKDDFFKSAMPBGYVOERTIFPK DDGNYKTRAEVIFEDTLVNRIELKIDIFKEDGNILKLELYNNSHNVYIMADKQX GIKYNFTRHNIEDGSVQLADHYQNTPIGDGPVLLPDNHYLSLTSQALSQDPNEKRDH MVLLFVTAAGITHTGMBELVYNGAIEODGLHAGSPAAWVERLFYDWAQTIIGCSDA AVFLSAGGRPVLVFKTDLSCALNELODEARLSWLATGTGVCAPVLDVVTAGRDWL LLGVPFGDILLSHLAPAEKVISIMADAMRLHTLDPATCPFDHQAQKRIERARTRMEA GLVDDDDLEDEGHQADIAELFARLKARMPDGEDLVVTHGDAELNINWENGRFSGFID CGLRGVADRYQDIALATRIAEELGGEWADRFVLVYGTAAQPSQRIAFYRLLDFF"
3455..3837
complement(4834..5493)
/gene="Cmr"
complement(4834..5493)

repeat_region
gene
complement(4834..5493)
/gene="Cmr"
complement(4834..5493)
```

```
/gene="Cmr"
/codon_start=1
/transl_table=11
/product="chloramphenicol acetyltransferase"
/protein_id="BAB32741.1"
/db_xref="GI:13094143"
/translacion="MEKINGTYTVDISQWHRKEHFEAFOSVACTNQTVHLDTAF LKTVKKKHFFYFAFIIHLARLNAHTEFRNAMKDGELVWDSVHPCYTVFHEQTETP SLSWSEYHDDFRQFLHYISQDVACGENLAYFPKGF TENMFFVSNPWVSTFSLNV ANMNFAPVFTMGKTYTQGDVKVLMPLAIQVHHAVCDGFHVGRLNELQQYCDSEWQGG A"

ORIGIN
Query Match 100.0%; Score 101; DB 12; Length 5590;
Best Local Similarity 100.0%; Pred. No. 2.4e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGCATGCTGGGATGCGGTGGGCTCTA 60
|
|
|
Db 925 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGCATGCTGGGATGCGGTGGGCTCTA 866
|
|
|
Qy 61 TGGCTTCTGAGCGGGAAGACCAAGCTGGGCTCTAGGGG 101
|
|
|
Db 865 TGGCTTCTGAGCGGGAAGACCAAGCTGGGCTCTAGGGG 825
|
|
|

RESULT 10
AY437643
LOCUS AY437643 5639 bp DNA circular SYN 10-NOV-2003
DEFINITION Expression vector pcGlobin 2, complete sequence.
ACCESSION AY437643
VERSION AY437643.1 GI:38155839
KEYWORDS Expression vector pcGlobin 2
SOURCE Expression vector pcGlobin 2
ORGANISM other sequences; artificial sequences; vectors.
REFERENCE 1 (bases 1 to 5639)
AUTHORS RO,H., Kim,E.J. and Rhee,M.
TITLE A new vector system, pcGlobin 2 for in vitro synthesized RNA injection into zebrafish embryos
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 5639)
AUTHORS RO,H., Kim,E.J. and Rhee,M.
TITLE Direct Submision
JOURNAL Submitted (14-OCT-2003) Department of Biology, College of Natural Sciences, Chungnam National University, 305-764, Daejeon 305-764, Korea

FEATURES
source
1. 5639
/organism="Expression vector pcGlobin 2"
/mol_type="other DNA"
/db_xref="taxon:254096"
/notes="eukaryotic expression vector for zebrafish embryo microinjection; derivative of pcDNA3"
complement(4643..5503)
/codon_start=1
/product="beta-lactamase"
/protein_id="AAR12689.1"
/db_xref="GI:38155840"
/translacion="MSIQHFRVALIPFFAAFLCLPVFAHPETLVKVKDAEDQLGARVGY IELDLSGKILSFRRPEERPFMMSTFKVLLCGAVLSRIDGAEQQLGRHHYQNQLDVE YSPVTRKHLTDGMTVRELCSAAITMSDNTAANLLLTITGIGKEITLAIHNNMGDVHTRL DRWEPELNEAIPNDRDITTPVAMATIRKLTLGELLTLASRQQLIDWMEADKRVAGPL LRSLAPAGWFIADKSGAGERSGIITAAALGPDGKPSRIVIVITTTGSOATMDERNQRIA EIGASLIKHW"

ORIGIN
Query Match 100.0%; Score 101; DB 12; Length 5639;
Best Local Similarity 100.0%; Pred. No. 2.4e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGCATGCTGGGATGCGGTGGGCTCTA 60
|
|
|
```

Db 1397 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGCATGCTGGGATCGCGTGGGCTCTA 1456
Qy 61 TGGCTTCTCAGCGGAAGAACACAGCTGGGGCTCTAGGGGG 101
Db 1457 TGGCTTCTCAGCGGAAGAACACAGCTGGGGCTCTAGGGGG 1497

RESULT 11
LOCUS AX2111282 5651 bp DNA linear PAT 06-SEP-2001
DEFINITION Sequence 6 from Patent WO0158493.
ACCESSION AX2111282
VERSION AX2111282.1 GI:15523691
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.

REFERENCE 1
AUTHORS Schambye,H.T., Andersen,K.V., van den Hazel,B., Christiansen,J. and Jeppesen,C.B.
TITLE Conjugates of follicle stimulating hormones
JOURNAL Patent: WO 0158493-A 6 16-AUG-2001;
Maxygen Aps (DK)
FEATURES Location/Qualifiers
source 1..5651
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
exon 1231..1617
/note="Coding sequence for human FSH-beta"

ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 5651;
Best Local Similarity 100.0%; Pred. No. 2.4e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGCATGCTGGGATCGCGTGGGCTCTA 60
Db 1826 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGCATGCTGGGATCGCGTGGGCTCTA 1885

Qy 61 TGGCTTCTCAGCGGAAGAACACAGCTGGGGCTCTAGGGGG 101
Db 1886 TGGCTTCTCAGCGGAAGAACACAGCTGGGGCTCTAGGGGG 1926

RESULT 12
LOCUS AX349366 5651 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 4 from Patent WO0202597.
ACCESSION AX349366
VERSION AX349366.1 GI:18615329
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.

REFERENCE 1
AUTHORS Okkels,J.S., Jensen,A.D. and van den Hazel,B.C.
TITLE Peptide extended glycosylated polypeptides
JOURNAL Patent: WO 0202597-A 4 10-JAN-2002;
Maxygen Aps (DK) ; Maxygen Holdings Ltd (KY)
FEATURES Location/Qualifiers
source 1..5651
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
exon 1231..1617
/note="Coding sequence for human FSH-beta"

ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 5651;
Best Local Similarity 100.0%; Pred. No. 2.4e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGCATGCTGGGATCGCGTGGGCTCTA 60
Db 1826 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGCATGCTGGGATCGCGTGGGCTCTA 1885

Qy 61 TGGCTTCTCAGCGGAAGAACACAGCTGGGGCTCTAGGGGG 101
Db 1886 TGGCTTCTCAGCGGAAGAACACAGCTGGGGCTCTAGGGGG 1926

RESULT 13
LOCUS AX202478 5731 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 66 from Patent WO0152620.
ACCESSION AX202478
VERSION AX202478.1 GI:15392206
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.

REFERENCE 1
AUTHORS Barbas,C.F., Stege,J.T., Guan,X. and Dalmia,B.
TITLE Methods and compositions to modulate expression in plants
JOURNAL Patent: WO 0152620-A 66 26-JUL-2001;
The Scripps Research Institute (US) ; SYNGENTA AGRICULTURAL DISCOVERY, INC. (CA)
FEATURES Location/Qualifiers
source 1..5731
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="2C7-SID"

ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 5731;
Best Local Similarity 100.0%; Pred. No. 2.4e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGCATGCTGGGATCGCGTGGGCTCTA 60
Db 1906 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGCATGCTGGGATCGCGTGGGCTCTA 1965

Qy 61 TGGCTTCTCAGCGGAAGAACACAGCTGGGGCTCTAGGGGG 101
Db 1966 TGGCTTCTCAGCGGAAGAACACAGCTGGGGCTCTAGGGGG 2006

RESULT 14
LOCUS AX685746 5995 bp DNA linear PAT 29-MAR-2003
DEFINITION Sequence 5 from Patent WO02102854.
ACCESSION AX685746
VERSION AX685746.1 GI:29371751
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.

REFERENCE 1
AUTHORS Thomassen-Wolf,E., Borges,E., Yayon,A. and Rom,E.
TITLE Antibodies that block receptor protein tyrosine Kinase activation, Methods of screening for and uses thereof
JOURNAL Patent: WO 02102854-A 5 27-DEC-2002;
MorphoSys AG (DE) ; ProChon Biotech Ltd. (IL)
FEATURES Location/Qualifiers
source 1..5995
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
exon <963..1670
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/db_xref="GI:29371752"

CDS

/translation=" DPEBPKSCDKTHTPCPAPELLGGPSVFLLPPKPKDTLMI SRT
PEVTCVVDVSHEDPEVKFNWYDGVGEVHNATKPRPEQYNSTRVVSVLTVLHQDWL
NGKEYCKGSNKALPAI EKTISKAGQPSQVYTLPPSRDELTKQVSLTCLVKGIF
YPSDIAVEWESNGQPPENYKTPPVLLDSGGSEFFLYSLKTVDKSRWQGNVPSSCVNHE
ALNHDIYTKSLSLSPG"

ORIGIN

	Query Match	100.0%;	Score 101;	DB 6;	Length 5995;
	Best Local Similarity	100.0%;	Pred. No. 2.4e-18;		
	Matches 101;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ACAGCAAGGGGAGGATTGGGAGACAAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTA	60		
Db	1870	ACAGCAAGGGGAGGATTGGGAGACAAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTA	1929		
Qy	61	TGGCTTCTAGGCGGGAAGAACCGAGCTGGGGCTCTAGGGG	101		
Db	1930	TGGCTTCTAGGCGGGAAGAACCGAGCTGGGGCTCTAGGGG	1970		

[illegible]

promoter	740. .887
	/gene="beta actin"
misc_feature	1734. .1776
	/note="multiple cloning site; MCS
	under control of CMV enhancer/chicken beta-actin promoter"
polyA_signal	1782. .1996
	/note="bovine growth hormone"
enhancer	2442. .2836
misc_feature	2908. .3723
	/note="eGFP
	under control of SV40 promoter"
polyA_signal	3757. .3887
	/note="late"
ORIGIN	

ORIGIN

	Query Match	100.0%;	Score 101;	DB 12;	Length 6084;
	Best Local Similarity	100.0%;	Pred. No. 2.4e-18;		
	Matches 101;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ACAGCAAGGGGGAGGATTGGGAACAAATAGCAGGCATGCTGGGGATGCGTGGGCTCTA	60		
Db	1951	ACAGCAAGGGGGAGGATTGGGAACAAATAGCAGGCATGCTGGGGATGCGTGGGCTCTA	60		
Qy	61	TGGCTTCTGAGGCGGGAAGAACCCAGCTGGGGCTCTAGGGGG	101		
Db	2011	TGGCTTCTGAGGCGGGAAGAACCCAGCTGGGGCTCTAGGGGG	2051		

Search completed: July 14, 2005, 14:03:30
Job time : 757.618 secs

GenCore version 5.1.6

Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:35:42 ; Search time 142.398 Seconds

(without alignments)
4198.742 Million cell updates/sec

Title: US-09-482-682-43_COPY_4141_4241

Perfect score: 101

Sequence: 1 acagcaagggggaggattgg.....ccagctggggctctaggggg 101

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002as:*

7: Geneseqn2002bs:*

8: Geneseqn2003as:*

9: Geneseqn2003bs:*

10: Geneseqn2003cs:*

11: Geneseqn2003ds:*

12: Geneseqn2004as:*

13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101	100.0	3482	2	ADH11353
2	101	100.0	4597	4	Aaf24901 Nucleotid
3	101	100.0	4825	13	ADRI12380
4	101	100.0	5015	10	ADB33528 Expressio
5	101	100.0	5082	2	ADH11417 Plasmid p
6	101	100.0	5218	12	ADM97811 & X UAS b
7	101	100.0	5302	12	ADI34681 Nucleotid
8	101	100.0	5425	2	ADH11233
9	101	100.0	5431	6	ABN86685 Nucleotid
10	101	100.0	5431	10	ADE21866 Plasmid v
11	101	100.0	5431	12	ADO5277 pcDNA3 pl
12	101	100.0	5432	3	AAZ89476 Transgeni
13	101	100.0	5446	2	AAV38297 Plasmid p
14	101	100.0	5446	6	AAI18619 Renilla l
15	101	100.0	5446	6	ABL53540 Vector pc
16	101	100.0	5446	12	ADN36314 Plasmid p
17	101	100.0	5458	6	ABL58494 Recombina
18	101	100.0	5458	6	ABL58493 Recombina
19	101	100.0	5614	6	ABL58489 Recombina
20	101	100.0	5614	6	ABL58490 Recombina

21	101	100.0	5651	5	AAI66195 Human FSH
22	101	100.0	5651	6	ABK40237 DNA encod
23	101	100.0	5695	6	ABL58492 Recombina
24	101	100.0	5695	6	ABL58491 Recombina
25	101	100.0	5695	8	ABT40262 PCXFc DNA
26	101	100.0	5695	8	ADA89054 Plasmid p
27	101	100.0	5695	10	ADG74306 Fibrobias
28	101	100.0	5731	4	AAD11615 Six fringe
29	101	100.0	5821	12	ADM97787 Gal4-DBD
30	101	100.0	5864	6	AAI44423 West nile
31	101	100.0	5864	6	AAI44424 West nile
32	101	100.0	6082	8	AAI56212 Human AB-
33	101	100.0	6082	8	AAI56211 Human AB-
34	101	100.0	6082	8	AAI56210 Human AB-
35	101	100.0	6085	8	AAI56213 Human AB-
36	101	100.0	6094	8	AAI56215 Human AB-
37	101	100.0	6097	8	AAI56214 Human AB-
38	101	100.0	6100	6	ABK96469 Plasmid p
39	101	100.0	6135	6	ABK96470 Plasmid p
40	101	100.0	6148	6	ABK15579 DNA encod
41	101	100.0	6149	6	ABK15580 DNA encod
42	101	100.0	6180	4	AAI13062 pcDNA3-B1
43	101	100.0	6186	5	AAI66194 Human FSH
44	101	100.0	6186	6	ABK40236 DNA encod
45	101	100.0	6195	6	ABK51585 Nuclear t

ALIGNMENTS

RESULT 1

ADH11353

ID ADH11353 standard; DNA; 3482 BP.

AC ADH11353;

DT 11-MAR-2004 (first entry)

DE Vertebrate UNC-53 protein homologue related nucleotide sequence.

KW UNC-53 vertebrate protein homologue; UNC-53; Caenorhabditis elegans;

KW cell shape regulator; cell motility regulator; cell migration;

KW cell behaviour regulator; phenotype; signal transduction pathway;

KW signal transducing protein; signal integrator protein;

KW neuronal regeneration; revascularisation; wound healing;

KW chronic neurodegenerative disease; acute traumatic injury;

KW fibrotic disease; gene; ds.

XX Unidentified.

XX OS

XX WO9824810-A2.

XX PD 11-JUN-1998.

XX PF 03-DEC-1997; 97WO-EP006956.

XX PR 04-DEC-1996; 96GB-00025283.

XX PA (JANC) JANSSEN PHARM NV.

XX PI Platteauw CJ, Buesa Arjol CM, Deraeymaeker M, Verhasselt P;

XX PI Pujol NJR, Maertens LJS, Luyten W, Geerts H, Vandekerckhove JS;

XX PI Geyssen J, Bogaert TAOE,

XX WPI; 1998-362411/31.

XX DR P-PSDB; ADH11354.

XX XX

XX PT Vertebrate homologues of C. elegans UNC-53 protein - useful for, e.g.

XX PT promoting neuronal regeneration, treating chronic neuro-degenerative

XX PT diseases-or acute traumatic injuries.

XX PS Disclosure; Page 414-417; 479pp; English.

XX XX

CC The present invention describes a vertebrate protein homologue of an UNC-
CC 53 protein of *Caenorhabditis elegans* or a functional equivalent,
CC derivative or bioprecursor of UNC-53. Also described: (1) a cDNA sequence
CC encoding a vertebrate homologue of the *C. elegans* UNC-53 protein; (2) a
CC nucleic acid which hybridises to the cDNA of (1); (3) vector comprising
CC the cDNA as in (1); (4) a host cell containing the vector as in (3); (5)
CC a transgenic cell, tissue or animal comprising the vector as in (3); (6)
CC a compound identified as an enhancer or inhibitor of the regulation of
CC cell shape, motility, or the direction of cell migration for use as a
CC therapeutic; (7) a method for determination of whether a protein is an
CC inhibitor or enhancer of regulation of cell behaviour, growth, shape or
CC motility or the direction of migration by contacting a host cell
CC expressing a homologue of UNC-53 and determining a change of phenotype;
CC (8) a method for identification of vertebrate homologues of *C. elegans*
CC UNC-53 gene by hybridising a probe of 15-50 bp of the unc-53 sequence to
CC a DNA library; and (9) a method for identification of a protein which is
CC active in the signal transduction pathway of a cell of which a vertebrate
CC homologue of UNC-53 is a component comprising: (i) contacting an extract
CC of a cell with an antibody to the UNC-53 homologue; (ii) identifying an
CC antibody/homologue complex; and (iii) analysing such a complex to
CC identify any non-antibody protein bound to the complex. UNC-53 is a
CC signal transducing or signal integrator protein involved in controlling
CC directionality of cell migration and cell shape in *C. elegans*. Vertebrate
CC homologues of UNC-53 can be used to promote neuronal regeneration,
CC revascularisation or wound healing, to treat chronic neurodegenerative
CC diseases or acute traumatic injuries or fibrotic diseases. The present
CC sequence is used in the exemplification of the present invention.
SQ Sequence 3482 BP; 767 A; 956 C; 913 G; 846 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 2; Length 3482;
Best Local Similarity 100.0%; Pred. No. 1.6e-22;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAGCAAGCGGAGGATTGGGAACAATAGCAGGCATGCTGGGATGCGGTGGGCTCTA 60
Db 1274 ACAGCAAGCGGAGGATTGGGAACAATAGCAGGCATGCTGGGATGCGGTGGGCTCTA 1333

Qy 61 TGGCTTCTGAGCGGAAAGAACCCAGCTGGGGCTCTAGGGGG 101
Db 1334 TGGCTTCTGAGCGGAAAGAACCCAGCTGGGGCTCTAGGGGG 1374

RESULT 2
AA24901
ID AAF24901 standard; DNA; 4597 BP.
XX
AC AAF24901;
XX
DT 20-APR-2001 (first entry)
XX
DE Nucleotide sequence of the plasmid pCDNA3.1/GS.
XX
KW Microsphere; dihydrazide; hyaluronic acid; inflammatory response;
KW myocardial ischemia; cardiac angiogenesis; haemophilia;
KW vascular endothelial growth factor; VEGF; ss.
XX
OS Synthetic.
XX
PN WO200078358-A2.
XX
PD 28-DEC-2000.
XX
PF 19-JUN-2000; 2000WO-US016837.
XX
PR 18-JUN-1999; 99US-0140260P.
XX
PA (COLL-) COLLABORATIVE GROUP LTD.
XX
PI Chen W;
XX
DR WPI; 2001-071363/08.

PT Hyaluronic acid micro spheres for use in gene therapy of myocardial
PT ischemia and hemophilia, comprising dihydrazide derivatized hyaluronic
XX acids crosslinked to nucleic acids.
XX
PS Example 1; Page 36-38; 38pp; English.
XX
CC The specification describes a microsphere comprising dihydrazide
CC derivatised hyaluronic acid crosslinked to a nucleic acid (NA). The
CC microspheres cause reduced inflammatory responses, and have increased
CC safety and biodegradability. The microspheres are useful for transfecting
CC a cell of a subject and for treating a subject having myocardial
CC ischemia by increasing cardiac angiogenesis. They are also useful for
CC treating haemophilia. The present sequence represents the plasmid
CC pCDNA3.1/GS, into which is inserted a polynucleotide sequence which is
CC crosslinked to hyaluronic acid. The polynucleotide sequence encodes a
CC vascular endothelial growth factor (VEGF)
XX
SQ Sequence 4597 BP; 1062 A; 1214 C; 1206 G; 1115 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 4; Length 4597;
Best Local Similarity 100.0%; Pred. No. 1.7e-22;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAGCAAGCGGAGGATTGGGAACAATAGCAGGCATGCTGGGATGCGGTGGGCTCTA 60
Db 1780 ACAGCAAGCGGAGGATTGGGAACAATAGCAGGCATGCTGGGATGCGGTGGGCTCTA 1839

Qy 61 TGGCTTCTGAGCGGAAAGAACCCAGCTGGGGCTCTAGGGGG 101
Db 1840 TGGCTTCTGAGCGGAAAGAACCCAGCTGGGGCTCTAGGGGG 1880

RESULT 3
ADRI2380
ID ADRI2380 standard; DNA; 4825 BP.
XX
AC ADRI2380;
XX
DT 21-OCT-2004 (first entry)
XX
DE Vector pMCP1.
XX
KW ss; cytostatic; VEGF modulator; angiogenesis inhibitor;
KW UTR-dependent expression; vascular endothelial growth factor;
KW untranslated region; cancer; angiogenesis; vector.
XX
OS Synthetic.
XX
PN WO2004065561-A2.
XX
PD 05-AUG-2004.
XX
PF 21-JAN-2004; 2004WO-US001643.
XX
PR 21-JAN-2003; 2003US-0441637P.
XX
PA (PTCT-) PTC THERAPEUTICS INC.
XX
PI Cao L, Trifillis P;
XX
DR WPI; 2004-571681/55.
XX

Identifying modulators of untranslated region-dependent expression of a
VEGF gene, useful for treating cancer, comprises contacting a compound
with a cell or translation mixture containing a reporter gene linked to a
VEGF gene UTR.

Example; SEQ ID NO 94; 251pp; English.

A method of identifying (M1) a compound that modulates untranslated
region-dependent expression of a vascular endothelial growth factor
(VEGF) gene comprises contacting a member of a library of compounds with
a cell or cell-free translation mixture containing a reporter gene

CC operably linked to an untranslated region (UTR) of the VEGF gene, and
 CC detecting expression of the reporter gene. A compound is identified as
 CC modulator if the level of expression of the reporter gene in the presence
 CC of the compound is altered as compared to that in the absence of the
 CC compound or in the presence of a control. Compounds identified by M1 are
 CC useful for treating, preventing or ameliorating cancer or its symptoms,
 CC and/or for inhibiting angiogenesis. This sequence corresponds to the
 CC vector pMCP1, a mammalian expression vector designed to integrate into
 CC the genome at sites containing the PRT recombination site using the flp
 CC recombinase.

XX SQ Sequence 4825 BP; 1236 A; 1135 C; 1204 G; 1250 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 13; Length 4825;
 Best Local Similarity 100.0%; Pred. No. 1.8e-22;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACACGAGGGGAGGATTGGGAACACATACACGAGCATCTGGGGATCGCGTGGGCTCTA 60
 |||||
 Db 2601 ACACGAGGGGAGGATTGGGAACACATACACGAGCATCTGGGGATCGCGTGGGCTCTA 2660
 Qy 61 TGGCTTCTGAGCGGGAAGAACACGAGCTGGGGCTCTAGGGGG 101
 |||||
 Db 2661 TGGCTTCTGAGCGGGAAGAACACGAGCTGGGGCTCTAGGGGG 2701

RESULT 4

ADB33528
 ID ADB33528 standard; DNA; 5015 BP.

XX AC ADB33528;

XX DT 04-DEC-2003 (first entry)

XX DE Expression vector nucleotide sequence SEQ ID NO:27.

XX KW fusion protein; amyloid precursor protein; APP; transcription factor;
 KW neurotropic; neuroprotective; APP inhibitor;
 KW amyloid precursor protein inhibitor; Alzheimer's disease; beta-secretase;
 KW gamma-secretase; human; gene; ds.

XX OS Synthetic.

XX OS Homo sapiens.

XX PN WO2003072041-A2.

XX PD 04-SEP-2003.

XX PF 23-FEB-2003; 2003WO-US005458.

XX PR 27-FEB-2002; 2002US-0360274P.

XX PA (MERI) MERCK & CO INC.

XX ES Espeseth AS, Ferrer M, Flores OA, Hazuda DJ, Inglese J;
 PI Miller MD, Register B, Shi X, Simon AJ, Zuck PD;

XX DR WPI; 2003-689968/65.

XX PT DNA encoding a fusion protein of amyloid precursor protein, useful in
 PT screening for anti-Alzheimer agents, comprises a fused transcription
 PT factor.

XX PS Disclosure; Fig 32B-F; 193pp; English.

XX CC The present invention describes a DNA molecule (I) that encodes a fusion
 CC protein (PP) comprising: (i) an amino acid sequence of amyloid precursor
 CC protein (APP), either the wild type, Swedish or NFEV versions; and (ii) a
 CC transcription factor (TF), fused in frame to the C-terminus of (i). Also
 CC described: (1) an expression vector containing (I); (2) a eukaryotic cell
 CC containing (I); and (3) methods for identifying a compound (A) that
 CC inhibits processing of APP, using the cells of (2). (I) has neurotropic and
 CC neuroprotective activities. (I) can be used to produce eukaryotic cells

CC that express PP and are useful in screening for agents that inhibit
 CC processing of APP. The agents are potentially useful for the treatment or
 CC prevention of Alzheimer's disease. Cells that express PP can screen for
 CC inhibitors of: (a) beta- and gamma-secretases; and (b)
 CC cytoplasmic/extracellular APP signaling in a single assay. Cell-based
 CC assays may be free of interference from alpha-secretase activity and are
 CC homogeneous (no chromatography, immunoprecipitation or washing required)
 CC so well suited to high-throughput screening. The present sequence
 CC represents a plasmid nucleotide sequence from the present invention.

XX SQ Sequence 5015 BP; 1167 A; 1297 C; 1279 G; 1272 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 10; Length 5015;
 Best Local Similarity 100.0%; Pred. No. 1.8e-22;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACACGAGGGGAGGATTGGGAACACATACGAGCATCTGGGGATCGCGTGGGCTCTA 60
 |||||
 Db 1190 ACACGAGGGGAGGATTGGGAACACATACGAGCATCTGGGGATCGCGTGGGCTCTA 1249
 Qy 61 TGGCTTCTGAGCGGGAAGAACACGAGCTGGGGCTCTAGGGGG 101
 |||||
 Db 1250 TGGCTTCTGAGCGGGAAGAACACGAGCTGGGGCTCTAGGGGG 1290

RESULT 5

ADH11417
 ID ADH11417 standard; DNA; 5082 BP.

XX AC ADH11417;

XX DT 11-MAR-2004 (first entry)

XX DE Plasmid pCB201 nucleotide sequence.

XX KW UNC-53 vertebrate protein homologue; UNC-53; Caenorhabditis elegans;
 KW cell shape regulator; cell motility regulator; cell migration;
 KW cell behaviour regulator; phenotype; signal transduction pathway;
 KW signal transducing protein; signal integrator protein;
 KW neuronal regeneration; revascularisation; wound healing;
 KW chronic neurodegenerative disease; acute traumatic injury;
 KW fibrotic disease; human; gene; ds.

XX OS Synthetic.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers
 FT CDS 1028..2149
 FT /*tag= a

XX PN WO9824810-A2.

XX PD 11-JUN-1998.

XX PF 03-DEC-1997; 97WO-EF006956.

XX PR 04-DEC-1996; 96GB-00025283.

XX PA (JANC) JANSSEN PHARM NV.

XX PI Platteeuw CJ, Buesa Arjol CM, Deraeymaeker M, Verhasselt P;

PI Pujol NJR, Maertens LJS, Luyten W, Geerts H, Vandekerckhove JS;

PI Geysen J, Bogaert TAOB;

XX DR WPI; 1998-362411/31.

XX DR P-PSDB; ADH11424.

XX PT Vertebrate homologues of C. elegans UNC-53 protein - useful for, e.g.
 PT promoting neuronal regeneration, treating chronic neuro-degenerative
 PT diseases or acute traumatic injuries.

XX PS Claim 96; SEQ ID NO 10; 479pp; English.

CC The present invention describes a vertebrate protein homologue of an UNC-
 CC 53 protein of *Caenorhabditis elegans* or a functional equivalent,
 CC derivative or bioprecursor of UNC-53. Also described: (1) a cDNA sequence
 CC encoding a vertebrate homologue of the *C. elegans* UNC-53 protein; (2) a
 CC nucleic acid which hybridizes to the cDNA of (1); (3) vector comprising
 CC the cDNA as in (1); (4) a host cell containing the vector as in (3); (5)
 CC a transgenic cell, tissue or animal comprising the vector as in (3); (6)
 CC a compound identified as an enhancer or inhibitor of the regulation of
 CC cell shape, motility, or the direction of cell migration for use as a
 CC therapeutic; (7) a method for determination of whether a protein is an
 CC inhibitor or enhancer of regulation of cell behaviour, growth, shape or
 CC motility or the direction of migration by contacting a host cell
 CC expressing a homologue of UNC-53 and determining a change of phenotype;
 CC (8) a method for identification of vertebrate homologues of *C. elegans*
 CC UNC-53 gene by hybridising a probe of 15-50 bp of the unc-53 sequence to
 CC a DNA library; and (9) a method for identification of a protein which is
 CC active in the signal transduction pathway of a cell of which a vertebrate
 CC homologue of UNC-53 is a component comprising: (i) contacting an extract
 CC of a cell with an antibody to the UNC-53 homologue; (ii) identifying an
 CC antibody/homologue complex; and (iii) analysing such a complex to
 CC identify any non-antibody protein bound to the complex. UNC-53 is a
 CC signal transducing or signal integrator protein involved in controlling
 CC directionality of cell migration and cell shape in *C. elegans*. Vertebrate
 CC homologues of UNC-53 can be used to promote neuronal regeneration,
 CC revascularisation or wound healing, to treat chronic neurodegenerative
 CC diseases or acute traumatic injuries or fibrotic diseases. The present
 CC sequence is used in the exemplification of the present invention.

XX SQ Sequence 5082 BP; 1164 A; 1365 C; 1311 G; 1242 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 2; Length 5082;
 Best Local Similarity 100.0%; Pred. No. 1.8e-22;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAGCAAGCGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATCGCGTGGGCTCTA 60
 Db 2874 ACAGCAAGCGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATCGCGTGGGCTCTA 2933

Qy 61 TGGCTTCTGAGCGGGAAGAACCAAGCTGGGGCTCTAGGGGG 101
 Db 2934 TGGCTTCTGAGCGGGAAGAACCAAGCTGGGGCTCTAGGGGG 2974

RESULT 6
 ADM97811
 ID ADM97811 standard; DNA; 5218 BP.
 XX AC ADM97811;
 XX DT 01-JUL-2004 (first entry)
 XX DE & X UAS beta-lactamase vector SEQ ID NO: 64.
 XX KW enzyme; sensor cell; signal transduction detection system; promoter;
 KW targeting sequence; targeted drug; ds; vector.
 XX OS Synthetic.
 XX OS Unidentified.
 XX PN WO2004031415-A2.
 XX PD 15-APR-2004.
 XX PF 05-SEP-2003; 2003WO-US028078.
 XX PR 05-SEP-2002; 2002US-0408297P.
 XX PA (VERT-) VERTEX PHARM INC.
 XX PI Whitney MA, Zeh K, Sanders PS;
 XX DR WPI; 2004-330208/30.

PT Developing a sensor cell, useful in determining the activity of a target
 PT gene and in developing therapeutic drugs, comprises providing cells
 PT comprising a signal transduction detection system and introducing DNA
 PT construct into cells.

XX PS Example 7; Page 231-234; 234pp; English.

XX CC The present invention relates to a method of developing a sensor cell,
 CC for determining the activity of a target gene in the cell, which
 CC comprises providing a homogeneous population of cells, where each of the
 CC cells comprises a signal transduction detection system and introducing
 CC into the population of cells an isolated DNA construct comprising a
 CC promoter operatively linked to a targeting sequence. The method is useful
 CC in developing a sensor cell for determining the activity of a target gene
 CC in the cell. The sensor cell and the methods are useful in developing new
 CC and therapeutic drugs directed to the targets. The present sequence is a
 CC vector used in the exemplification of the invention.

XX SQ Sequence 5218 BP; 1231 A; 1361 C; 1335 G; 1291 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 12; Length 5218;
 Best Local Similarity 100.0%; Pred. No. 1.8e-22;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACAGCAAGCGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATCGCGTGGGCTCTA 60
 Db 4177 ACAGCAAGCGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATCGCGTGGGCTCTA 4236

Qy 61 TGGCTTCTGAGCGGGAAGAACCAAGCTGGGGCTCTAGGGGG 101
 Db 4237 TGGCTTCTGAGCGGGAAGAACCAAGCTGGGGCTCTAGGGGG 4277

RESULT 7
 ADI34681
 ID ADI34681 standard; DNA; 5302 BP.
 XX AC ADI34681;
 XX DT 22-APR-2004 (first entry)
 XX DE Nucleotide sequence of plasmid pcDNA6/Biotag (TM) -D-TOPO.
 XX KW Recombinational cloning; recombination; topoisomerase;
 KW fusion protein production; ds.
 XX OS Synthetic.
 XX PN WO2004005482-A2.
 XX PD 15-JAN-2004.
 XX PF 08-JUL-2003; 2003WO-US021339.
 XX PR 08-JUL-2002; 2002US-0393756P.
 XX PR 19-JUL-2002; 2002US-0396627P.
 XX PR 10-OCT-2002; 2002US-0417172P.
 XX PA (INVI-) INVITROGEN CORP.
 XX PI Bennett RP;
 XX DR WPI; 2004-091356/09.
 XX PT New isolated nucleic acid molecules having one or more recombination
 PT sites and encoding an amino acid sequence tag, useful for recombination
 PT and/or topoisomerase-mediated cloning methods for producing fusion
 PT proteins.
 XX Example 1; Fig 11A-B; 135pp; English.
 XX CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC one or more recombination sites, and one or more nucleic acid sequences

CC which encode an amino acid sequence tag. (1) can also comprise one or
CC more topoisomerase recognition sites and/or one or more topoisomerases.
CC The amino acid sequence tag is an amino acid sequence that is capable of
CC being post-translationally modified, and is an amino acid sequence that
CC is capable of being post-translationally modified by biotinylation,
CC attachment of 4-phosphopantetheine, attachment of lipoid acid or
CC attachment of flavins, and is an amino acid sequence that is capable of
CC being biotinylated, wherein the amino acid sequence that is capable of
CC being biotinylated is all or a portion of the Klebsiella pneumoniae
CC oxalacetate decarboxylase a subunit, all or a portion of the
CC Propionibacterium shermanii transcarboxylase 1.3S subunit, or all or a
CC portion of the Escherichia coli biotin carboxyl carrier protein component
CC of acetyl-CoA carboxylase. The methods and compositions of the present
CC invention are useful for identifying, concentrating, purifying and/or
CC producing fusion proteins that comprise an amino acid sequence tag. The
CC nucleic acid molecules can also be used in recombinational cloning and/or
CC topoisomerase-mediated cloning methods to produce polynucleotide
CC constructs which encode the fusion proteins. The present sequence
CC represents the nucleotide sequence of a plasmid pcDNA6/Biotag(TM)-D-TOPO
XX Sequence 5302 BP; 1254 A; 1361 C; 1349 G; 1338 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 12; Length 5302;
Best Local Similarity 100.0%; Pred. No. 1.8e-22;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATCGGTGGGCTCTTA 60
Db 1425 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATCGGTGGGCTCTTA 1484
QY 61 TGGCTTCTGAGCGGAAAGAACACAGCTGGGGCTCTAGGGGG 101
Db 1485 TGGCTTCTGAGCGGAAAGAACACAGCTGGGGCTCTAGGGGG 1525

RESULT 8
ADH11233
ID ADH11233 standard; DNA; 5425 BP.
AC ADH11233;
DT 11-MAR-2004 (first entry)
XX Vertebrate UNC-53 protein homologue related nucleotide sequence.
DE UNC-53 vertebrate protein homologue; UNC-53; Caenorhabditis elegans;
KW cell shape regulator; cell motility regulator; cell migration;
KW cell behaviour regulator; phenotype; signal transduction pathway;
KW signal transducing protein; signal integrator protein;
KW neuronal regeneration; revascularisation; wound healing;
KW chronic neurodegenerative disease; acute traumatic injury;
KW fibrotic disease; gene; ds.
XX Unidentified.
OS WO9824810-A2.
PN 11-JUN-1998.
PD 03-DEC-1997; 97WO-EP006956.
PF 04-DEC-1996; 96GB-00025283.
PR (JANC) JANSSEN PHARM NV.
PA Platteeuw CJ, Buesa Arjol CM, Deraeymacker M, Verhasselt P;
PI Puijol NJR, Maertens LJS, Luyten W, Geerts H, Vandekerckhove JS;
PI Geysen J, Bogaert TAOE;
XX WPI; 1998-362411/31.
DR P-PSDB; ADH11234.
XX Vertebrate homologues of C. elegans UNC-53 protein - useful for, e.g.

PT promoting neuronal regeneration, treating chronic neuro-degenerative
XX diseases or acute traumatic injuries.
PS Disclosure; Page 231-237; 479pp; English.
XX The present invention describes a vertebrate protein homologue of an UNC-
CC 53 protein of Caenorhabditis elegans or a functional equivalent,
CC derivative or bioprecursor of UNC-53. Also described: (1) a cDNA sequence
CC encoding a vertebrate homologue of the C. elegans UNC-53 protein; (2) a
CC nucleic acid which hybridises to the cDNA of (1); (3) vector comprising
CC the cDNA as in (1); (4) a host cell containing the vector as in (3); (5)
CC a transgenic cell, tissue or animal comprising the vector as in (3); (6)
CC a compound identified as an enhancer or inhibitor of the regulation of
CC cell shape, motility, or the direction of cell migration for use as a
CC therapeutic; (7) a method for determination of whether a protein is an
CC inhibitor or enhancer of regulation of cell behaviour, growth, shape or
CC motility or the direction of migration by contacting a host cell
CC expressing a homologue of UNC-53 and determining a change of phenotype;
CC (8) a method for identification of vertebrate homologues of C. elegans
CC unc-53 gene by hybridising a probe of 15-50 bp of the unc-53 sequence to
CC a DNA library; and (9) a method for identification of a protein which is
CC active in the signal transduction pathway of a cell of which a vertebrate
CC homologue of UNC-53 is a component comprising: (i) contacting an extract
CC of a cell with an antibody to the UNC-53 homologue; (ii) identifying an
CC antibody/homologue complex; and (iii) analysing such a complex to
CC identify any non-antibody protein bound to the complex. UNC-53 is a
CC signal transducing or signal integrator protein involved in controlling
CC directionalities of cell migration and cell shape in C. elegans. Vertebrate
CC homologues of UNC-53 can be used to promote neuronal regeneration,
CC revascularisation or wound healing, to treat chronic neurodegenerative
CC diseases or acute traumatic injuries or fibrotic diseases. The present
CC sequence is used in the exemplification of the present invention.
XX
SQ Sequence 5425 BP; 1250 A; 1463 C; 1420 G; 1292 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 2; Length 5425;
Best Local Similarity 100.0%; Pred. No. 1.8e-22;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATCGGTGGGCTCTTA 60
Db 3217 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATCGGTGGGCTCTTA 3276
QY 61 TGGCTTCTGAGCGGAAAGAACACAGCTGGGGCTCTAGGGGG 101
Db 3277 TGGCTTCTGAGCGGAAAGAACACAGCTGGGGCTCTAGGGGG 3317
RESULT 9
ABN86685
ID ABN86685 standard; DNA; 5431 BP.
XX AC ABN86685;
XX 05-NOV-2002 (first entry)
DT Nucleotide sequence of a pcDNA3 plasmid vector.
DE Major histocompatibility complex; MHC; antigen presenting cell; APC;
XX antigen; cytostatic; virucide; gene therapy; CD8; vaccine; therapeutic;
KW cancer; viral infection; ds.
XX Synthetic.
OS WO200261113-A2.
PN 08-AUG-2002.
PD 01-FEB-2002; 2002WO-US002598.
PF 01-FEB-2001; 2001US-0265334P.
PR (UVJO) UNIV JOHNS HOPKINS.

XX Wu T, Hung C;
XX WPI; 2002-619261/66.
XX
XX Nucleic acid molecule encoding a fusion polypeptide that promotes
XX processing via the Major Histocompatibility Complex class I pathway
XX and/or promotes activity of an antigen presenting cell, useful as vaccine
XX for cancer and viral infections.
XX
XX Claim 24; Page 22-23; 127pp; English.
XX
XX The invention relates to a new nucleic acid molecule (I) encoding a
XX fusion polypeptide useful as a vaccine composition. (I) comprises a first
XX nucleic acid sequence encoding a first polypeptide or peptide that
XX promotes processing via the Major Histocompatibility Complex (MHC) class
XX I pathway (MHC-I-PP) and/or promotes development or activity of an
XX antigen presenting cell (APC). The second nucleic acid sequence is linked
XX in frame to the first nucleic acid sequence or to a linker nucleic acid
XX sequence and encodes an antigenic polypeptide or peptide. The methods and
XX compositions of the present invention are useful as therapeutic vaccine
XX for cancer and for major viral infections, such as hepatoma and cervical
XX cancer, that cause morbidity and mortality. They can also be used in
XX treating animal diseases, such as equine herpesvirus, bovine viruses,
XX Marek's disease, retroviral and lentiviral diseases and rabies, in the
XX veterinary medicine context. The present sequence represents the
XX nucleotide sequence of a pcDNA3 plasmid vector
XX
SQ Sequence 5431 BP; 1253 A; 1413 C; 1386 G; 1379 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 6; Length 5431;
Best Local Similarity 100.0%; Pred. No. 1.8e-22;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATCGCGTGGGCTCTA 60
Db 1189 ACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATCGCGTGGGCTCTA 1248
Qy 61 TGGCTTCTGAGCGGAAAGAACCACTGGGCTCTAGGGGG 101
Db 1249 TGGCTTCTGAGCGGAAAGAACCACTGGGCTCTAGGGGG 1289
RESULT 10
ADE21866
ID ADE21866 standard; DNA; 5431 BP.
XX
XX ADE21866;
XX
XX 29-JAN-2004 (first entry)
XX
XX Plasmid vector pcDNA3 nucleotide sequence SEQ ID NO:8.
XX
XX chimeric fusion; translocation; antigenic; cytostatic; immunotherapy;
XX gene therapy; cancer; tumour; gene; ds.
XX
XX Synthetic.
XX
XX WO2003085085-A2.
XX
XX 16-OCT-2003.
XX
XX 04-APR-2003; 2003WO-US010235.
XX
XX 04-APR-2002; 2002US-00115440.
XX
XX (UYJO) UNIV JOHNS HOPKINS.
XX
XX Wu T, Hung C;
XX WPI; 2003-877027/81.
XX
XX New nucleic acid encoding a chimeric fusion or fusion polypeptide

PT comprising a first domain with a translocation polypeptide, and a second
PT domain with an antigen having at least one antigenic peptide, useful for
PT treating cancer.
XX
XX Disclosure; SEQ ID NO 8; 68pp; English.
XX
XX The present invention describes a nucleic acid (I) encoding a chimeric
XX fusion or fusion polypeptide comprising a first domain with a
XX translocation polypeptide, and a second domain comprising an antigen
XX having at least one antigenic peptide. Also described: (1) an expression
XX vector comprising (I) operatively linked to a promoter and optionally, to
XX one or more regulatory elements that enhance expression of the nucleic
XX acid in a cell; (2) a particle comprising (I) or the expression vector;
XX (3) a cell that has been modified to comprise (I) or the expression
XX vector; (4) a chimeric polypeptide comprising a first domain with a
XX translocation polypeptide, and a second domain comprising an antigen
XX having at least one antigenic peptide; (5) a pharmaceutical composition
XX capable of inducing or enhancing an antigen specific immune response,
XX comprising (I), expression vector, particle, cell, cell of the particle,
XX or the chimeric polypeptide; and a carrier or excipient; (6) inducing or
XX enhancing an antigen specific immune response by administering the
XX composition described above; (7) increasing the number of CD8 + CTLs
XX specific for a selected desired antigen in a subject by administering the
XX composition described above; and (8) inhibiting the growth of a tumour in
XX a subject by administering the composition described above. (I) has
XX cytostatic activity, and can be used in immunotherapy, and gene therapy.
XX The nucleic acids (I), compositions and methods are useful for treating
XX cancer. The present sequence represents a plasmid vector nucleotide
XX sequence which is used in the exemplification of the present invention.
XX
SQ Sequence 5431 BP; 1253 A; 1413 C; 1386 G; 1379 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 10; Length 5431;
Best Local Similarity 100.0%; Pred. No. 1.8e-22;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATCGCGTGGGCTCTA 60
Db 1189 ACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATCGCGTGGGCTCTA 1248
Qy 61 TGGCTTCTGAGCGGAAAGAACCACTGGGCTCTAGGGGG 101
Db 1249 TGGCTTCTGAGCGGAAAGAACCACTGGGCTCTAGGGGG 1289
RESULT 11
ADO05277
ID ADO05277 standard; DNA; 5431 BP.
XX
XX ADO05277;
XX
XX 29-JUL-2004 (first entry)
XX
XX pcDNA3 plasmid vector.
XX
XX Translocation domain; bacterial toxin; exotoxin A domain II; ETA;
XX major histocompatibility complex; MHC class I; vaccine; immune response;
XX CD8+ cytotoxic T lymphocyte; CTL; tumour; E7 antigen; pcDNA3 plasmid; ds.
XX
XX Synthetic.
XX
XX US2004086845-A1.
XX
XX 06-MAY-2004.
XX
XX 04-APR-2002; 2002US-00115440.
XX
XX 20-OCT-1999; 99US-00421608.
XX
XX 09-FEB-2000; 2000US-00501097.
XX
XX 20-OCT-2000; 2000WO-US041422.
XX
XX 04-APR-2001; 2001US-0281003P.
XX
XX (WUTT/) WU T.

PA (HUNG/) HUNG C.
XX
PI Wu T, Hung C;
XX
DR WPI; 2004-356187/33.
XX
XX Novel chimeric polypeptide e.g., *Pseudomonas aeruginosa* exotoxin A domain
PT II/human papilloma virus-16 E7 peptide useful for inducing or enhancing
PT antigen specific immune response, or for inhibiting growth of tumor in
PT subject.
XX
XX Disclosure; SEQ ID NO 8; 48pp; English.
XX
XX The invention relates to nucleic acid encoding a chimeric polypeptide
CC comprising a translocation domain of a bacterial toxin and at least one
CC antigenic peptide. The preferred translocation domain is domain II of
CC *Pseudomonas aeruginosa* exotoxin A (ETA(dII)) and the preferred antigen is
CC human papilloma virus type 16 (HPV-16) E7 which is a model tumour
CC antigen. The antigenic peptide comprises an epitope that binds to and is
CC presented on the cell surface by major histocompatibility complex (MHC)
CC class I proteins. The nucleic acid of the invention is useful as vaccine
CC composition for enhancing antigen specific immune response, increasing
CC the number of CD8+ cytotoxic T lymphocytes (CTLs) and for inhibiting the
CC growth of a tumour. The present sequence is pcDNA3 plasmid vector used in
CC the invention.
XX
SQ Sequence 5431 BP; 1253 A; 1413 C; 1386 G; 1379 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 12; Length 5431;
Best Local Similarity 100.0%; Pred. No. 1.8e-22;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACAGCAAGGGGGAGGATTGGGAAGACAAATAGCAGGCATGCTGGGGATCGGTGGGCTCTA 60
DB 1189 ACAGCAAGGGGGAGGATTGGGAAGACAAATAGCAGGCATGCTGGGGATCGGTGGGCTCTA 1248
QY 61 TGGCTTCTGAGCGGAAAGAACACAGCTGGGGCTCTAGGGGG 101
DB 1249 TGGCTTCTGAGCGGAAAGAACACAGCTGGGGCTCTAGGGGG 1289
RESULT 12
AAZ89476
ID AAZ89476 standard; DNA; 5432 BP.
XX
AC AAZ89476;
XX
XX
DT 22-JUN-2000 (first entry)
XX
DE Transgenic APP DNA #2.
XX
XX APP; amyloid precursor protein; gamma-secretase; neuroprotective;
KW nontropic; transgenic; Alzheimer's disease; Down's syndrome; ds.
XX
XX Synthetic.
OS
XX DE19856261-C1.
PN
XX 30-MAR-2000.
PD
XX 07-DEC-1998; 98DE-01056261.
PF
XX 07-DEC-1998; 98DE-01056261.
PR
XX (HMR1) HOECHST MARION ROUSSEL DEUT GMBH.
PA
XX Peraus G;
PI
XX WPI; 2000-258119/23.
DR
XX
XX Detection of gamma-secretase by detection of A-beta peptide useful for
PT determining gamma-secretase activity and for identifying inhibitors.
PT
XX

PS Claim 30; Page 7-8; 16pp; German.
XX
XX This invention describes a novel method for the detection of human gamma-
CC secretase by detection of a partial protein formed by cleavage of a
CC fusion protein encoded by a transgene containing a first nucleotide
CC sequence which encodes a protein comprising the amino acid sequence (A)
CC and a second nucleotide sequence which encodes a signal peptide. The
CC products of the invention have neuroprotective and nontropic activity.
CC The method is used to detect activity of gamma-secretase. The transgene
CC and/or vectors are useful for the production of a transgenic cell or C.
CC *elegans*. Transgenic C. *elegans* is useful in a method for the
CC determination of gamma-secretase activity. The transgenic C. *elegans* is
CC also useful in a method to identify inhibitors of the gamma-secretase
CC activity. The methods and transgenes are useful in research of
CC Alzheimer's disease. Inhibitors of gamma-secretase are useful in
CC control/treatment of Alzheimer's and possibly Down's syndrome. This
CC sequence encodes a transgenic amyloid precursor protein (APP) which is
CC described in the method of the invention
XX
SQ Sequence 5432 BP; 1251 A; 1410 C; 1390 G; 1381 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 3; Length 5432;
Best Local Similarity 100.0%; Pred. No. 1.8e-22;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACAGCAAGGGGGAGGATTGGGAAGACAAATAGCAGGCATGCTGGGGATCGGTGGGCTCTA 60
DB 1190 ACAGCAAGGGGGAGGATTGGGAAGACAAATAGCAGGCATGCTGGGGATCGGTGGGCTCTA 1249
QY 61 TGGCTTCTGAGCGGAAAGAACACAGCTGGGGCTCTAGGGGG 101
DB 1250 TGGCTTCTGAGCGGAAAGAACACAGCTGGGGCTCTAGGGGG 1290
RESULT 13
AAV38297
ID AAV38297 standard; DNA; 5446 BP.
XX
AC AAV38297;
XX
XX
DT 17-OCT-2003 (revised)
DT 26-OCT-1998 (first entry)
XX
DE Plasmid pcDNA3.
XX
XX Plasmid pcDNA3; pneumococcal surface protein A; PspA; infection;
KW Streptococcus pneumoniae; sepsis; otitis media; meningitis; bacteraemia;
KW pneumonia; vaccine; genetic immunisation; ss.
XX
XX Human herpesvirus 5.
OS
XX Chimeric.
XX
XX WO9824927-A1.
PN
XX 11-JUN-1998.
XX
XX 04-DEC-1997; 97WO-US022847.
PF
XX 04-DEC-1996; 96US-00759505.
PR
XX (UYAL-) UNIV ALABAMA.
PA
XX Briles DE, McDaniel LS, Curriel DT;
PI
XX WPI; 1998-333343/29.
DR
XX Plasmid containing pneumococcal epitope for expression in eukaryotic
PT cells - useful for eliciting immunological response to pneumococcal
PT infection or sepsis.
XX
XX Example 1; Fig 1B1 to 1B-5; 47pp; English.
PS
XX This is the DNA sequence of plasmid pcDNA3 (Invitrogen). A portion of the
CC

CC gene (see AAV38298) that codes for respiratory syncytial virus
CC glycoprotein G (RSVG) has been amplified, digested with KpnI and ligated
CC into KpnI-digested pcDNA3 upstream of the multiple cloning site of pcDNA3
CC and downstream of the human cytomegalovirus immediate early (HCMV-IE)
CC promoter to create pGT41. A full-length coding sequence of Streptococcus
CC pneumoniae Rxi pneumococcal surface protein A (PspA) was then inserted
CC into pGT41 to create a fusion between RSVG and PspA. Intramuscular
CC immunisation of BALB/c mice with the resulting plasmid, designated
CC pKSD2601, induced protection against an otherwise lethal challenge with a
CC capsular type 3 pneumococcus. A plasmid for expression of pneumococcal
CC epitope DNA in eukaryotic cells is claimed. The plasmid includes a
CC promoter for driving expression in a eukaryotic cell (e.g. HCMV-IE), DNA
CC encoding a leader sequence (e.g. of RSVG) which facilitates expression,
CC translation through or transport of the expression product in a
CC eukaryotic cell membrane, and DNA encoding a pneumococcal epitope such as
CC PspA. The invention also provides a vaccine comprising the plasmid and a
CC suitable carrier or diluent, and optionally one or more cytokines or DNA
CC encoding them, or a bacterial delivery system. The vaccine is used to
CC elicit an immunological response in a host, including humans, susceptible
CC to pneumococcal infection or sepsis. The plasmid can also be used to
CC express a pneumococcal epitope of interest in vitro. (Updated on 17-OCT-
CC 2003 to standardise OS field)

XX SQ Sequence 5446 BP; 1255 A; 1417 C; 1390 G; 1384 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 2; Length 5446;

Best Local Similarity 100.0%; Pred. No. 1.8e-22;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGCAAGGGGGAGGATTGGGAGACAAATAGCAGCATGCTGGGGATCGCGTGGCTCTA 60

DB 1204 ACAGCAAGGGGGAGGATTGGGAGACAAATAGCAGCATGCTGGGGATCGCGTGGCTCTA 1263

QY 61 TGGCTTCTGAGCGGGAAGAACCACTGGGGCTCTAGGGGG 101

DB 1264 TGGCTTCTGAGCGGGAAGAACCACTGGGGCTCTAGGGGG 1304

RESULT 14

AAS18619

ID AAS18619 standard; DNA; 5446 BP.

XX AC AAS18619;

XX DT 26-FEB-2002 (first entry)

XX DE Renilla luciferase vector pcDNA3.

XX KW Renilla luciferase; sea pansy; cell proliferation disorder;
KW immune disorder; hypogammaglobulinaemia; haematologic condition; anaemia;
KW neoplasm; cancer; human immunodeficiency virus; HIV;
KW tissue white cell infiltrative disorder; organ failure;
KW myotrophic condition; gonadal failure; bone disorder; muscle disorder;
KW osteoporosis; endocrine condition; vascular disorder; atherogenesis;
KW pcDNA3; ds.

XX OS Synthetic.

XX PN WO200181614-A2.

XX PD 01-NOV-2001.

XX PF 25-APR-2001; 2001WO-US013512.

XX PR 25-APR-2000; 2000US-00559874.

XX PR 02-JUN-2000; 2000US-00586339.

XX PA (CHEM-) CHEMICON INT.

XX PI Leng J;

XX DR WPI; 2002-041420/05.

XX

PT Determining cell proliferation for monitoring treatment of a subject,
PT comprises obtaining light emission data from cell containing Renilla
PT luciferase for specific time, and detecting a change in the data.

XX Example 1; Fig 1A-B; 52pp; English.

PS The invention describes a novel method for measuring proliferation of a
XX cell or population of cells. The method comprises obtaining light
CC emission data from a cell containing a Renilla luciferase over a period
CC of time, cell proliferation of a cell or a population of cells can be
CC measured by a change in light emission data indicating proliferation. A
CC vector containing the Renilla luciferase enzyme is useful for diagnosing
CC a cell proliferative disorder including: neoplasm or cancer, viral
CC disorder or disease e.g. Human immunodeficiency virus (HIV), immune
CC disorders e.g. hypogammaglobulinaemia and haematologic conditions e.g.
CC anaemias, tissue white cell infiltrative disorders, organ failure,
CC myotrophic conditions, gonadal failure, conditions of bone and muscle
CC e.g. osteoporosis, endocrine conditions and vascular disorders e.g.
CC atherogenesis, by transfecting a cell obtained from a subject with the
CC vector and comparing the light emission data from the cell to that of a
CC cell which does not have a cell proliferative disorder. A difference in
CC light emission is indicative of a cell proliferative disorder. The vector
CC is also useful for determining the effect of an agent on cell
CC proliferation, by transfecting a cell obtained from a sample with the
CC vector, and contacting the transfected cell with an agent suspected of
CC modulating cell proliferation under conditions that allow the agent and
CC the cell to interact, and comparing the light emission data from the cell
CC to the light emission data from the cell in the absence of the agent. The
CC methods are useful for drug discovery and drug screening, and in
CC monitoring the treatment of a subject diagnosed with a cell proliferative
CC disorder. This sequence is the vector pcDNA3 into which Renilla
CC luciferase (AAS18616) is placed before transformation of cells with the
CC vector to allow measurement of cell proliferation described in the method
XX of the invention

SQ Sequence 5446 BP; 1255 A; 1417 C; 1390 G; 1384 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 6; Length 5446;

Best Local Similarity 100.0%; Pred. No. 1.8e-22;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACAGCAAGGGGGAGGATTGGGAGACAAATAGCAGCATGCTGGGGATCGCGTGGCTCTA 60

DB 1204 ACAGCAAGGGGGAGGATTGGGAGACAAATAGCAGCATGCTGGGGATCGCGTGGCTCTA 1263

QY 61 TGGCTTCTGAGCGGGAAGAACCACTGGGGCTCTAGGGGG 101

DB 1264 TGGCTTCTGAGCGGGAAGAACCACTGGGGCTCTAGGGGG 1304

RESULT 15

ABL53540

ID ABL53540 standard; DNA; 5446 BP.

XX AC ABL53540;

XX DT 10-JUN-2002 (first entry)

XX DE Vector pcDNA3.

XX KW Vector; pcDNA3; heat shock protein 60; Hsp60; autoimmune disease;
KW insulin dependent diabetes mellitus; IDDM; DNA immunisation; vaccine;
KW CpG; antidiabetic; immunotherapy; gene therapy; ds.

XX OS Cytomegalovirus.

XX OS Bos taurus.

XX OS Unidentified.

XX OS Chimeric.

XX PN WO200216549-A2.

XX PD 28-FEB-2002.

XX

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 .Compugen Ltd.
OM nucleic - nucleic search, using sw model
Run on: July 14, 2005, 05:15:57 ; Search time 961.667 Seconds
(without alignments)
3997.736 Million cell updates/sec
Title: US-09-482-682-43_COPY_4141_4241
Perfect score: 101
Sequence: 1 acagcaaggaggagattgg.....ccagctggggctctaggggg 101
Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 34239544 seqs, 19032134700 residues
Total number of hits satisfying chosen parameters: 68479088
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : EST:
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101	100.0	295	CN778129	pgn2c.pk0
2	85	84.2	378	CF315931	HD--05-A1
3	70.6	69.9	640	CN161285	950563 MA
4	66.4	65.7	333	EX993205	Reverse s
5	66	65.3	772	BZ851581	CH240 281
6	62.6	62.0	755	CF462023	932791 MA
7	59.4	58.8	683	CF366325	840970 MA
8	58.6	58.0	648	CB419493	592356 MA
9	54.4	53.9	666	CR010518	Forward s
10	49.2	48.7	754	CR046705	Forward s
11	48.2	47.7	119	CR159589	Forward s
12	48.2	47.7	422	CB763949	Forward s
13	48.2	47.7	433	CB760260	AMGNNUC:M
14	48.2	47.7	434	CB759281	AMGNNUC:M
15	48.2	47.7	449	CB742761	AMGNNUC:M
16	48.2	47.7	449	CB742762	AMGNNUC:M
17	48.2	47.7	450	CB788058	AMGNNUC:M
18	48.2	47.7	452	CB786684	AMGNNUC:M
19	48.2	47.7	457	CB740095	AMGNNUC:M
20	48.2	47.7	464	CB734940	AMGNNUC:M
21	48.2	47.7	467	CB714287	AMGNNUC:M
22	48.2	47.7	470	CB733007	AMGNNUC:M
23	48.2	47.7	472	CB730485	AMGNNUC:M
24	48.2	47.7	475	CB729030	AMGNNUC:M

25	48.2	47.7	482	6	CB728068	AMGNNUC:M
26	48.2	47.7	482	6	CB728069	AMGNNUC:M
27	48.2	47.7	482	6	CB728070	AMGNNUC:M
28	48.2	47.7	483	6	CB727859	AMGNNUC:M
29	48.2	47.7	484	6	CB727358	AMGNNUC:M
30	48.2	47.7	487	6	CB726627	AMGNNUC:M
31	48.2	47.7	487	6	CB726628	AMGNNUC:M
32	48.2	47.7	488	6	CB726362	AMGNNUC:M
33	48.2	47.7	496	6	CB713072	AMGNNUC:M
34	47.2	46.7	448	6	CB746136	AMGNNUC:M
35	47.2	46.7	450	6	CB788057	AMGNNUC:M
36	47.2	46.7	455	6	CB741728	AMGNNUC:M
37	46.2	45.7	439	6	CB750081	AMGNNUC:M
38	45	44.6	723	7	CK468080	939446 MA
c 39	44.6	44.2	207	9	CR162425	Forward s
c 40	43.4	43.0	798	7	CN158304	947089 MA
c 41	43.2	42.8	733	9	CR000527	Forward s
42	42.8	42.4	699	7	CN164031	994145 MA
43	42.8	42.4	778	7	CN161630	950941 MA
44	40.2	39.8	667	6	CB423182	596384 MA
45	39.6	39.2	752	7	CN164455	994805 MA

ALIGNMENTS

RESULT 1
CN778129
LOCUS
DEFINITION
pgn2c.pk001.h10.f Chicken Lymphoid cDNA library (pgn2c) Gallus
gallus sequence.
295 bp mRNA linear EST 20-MAY-2004
CN778129.1 GI:47548763
ACCESSION
VERSION
EST
KEYWORDS
SOURCE
ORGANISM
Gallus gallus (chicken)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE
1 (bases 1 to 295)
Morgan, R.W. and Burnside, J.
Chicken ESTs from lymphoid tissue- 3' sequence
Unpublished (2004)
JOURNAL
COMMENT
Contact: Robin W. Morgan
University of Delaware
Townsend Hall, Newark, DE 19717, USA
Tel: 302-831-1341
Fax: 302-831-2822
Email: morgan@udel.edu, www.chickest.udel.edu.
FEATURES
Location/Qualifiers
source
1..295
/organism="Gallus gallus"
/mol_type="mRNA"
/db_xref="taxon:9031"
/clones="pgn2c.pk001.h10.f 3'end of pat.pk0008.d12"
/sex="Male and Female"
/tissue_type="thymus, bursa, spleen, PBL, bone marrow"
/lab_host="E.Coli EMDH10B"
/clone_lib="Chicken Lymphoid cDNA library (pgn2c)"
/note="Vector: PCWVSPORT 6"

Query Match	100.0%	Score	101;	DB	7;	Length	295;
Best Local Similarity	100.0%	Pred. No.	9.8e-20;				
Matches	101;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	ACAGCAAGGGGAGGATTGGGAAGACATACAGCATGCTGGGATCGCGTGGGCTCTA	60				
Db	188	ACAGCAAGGGGAGGATTGGGAAGACATACAGCATGCTGGGATCGCGTGGGCTCTA	247				
Qy	61	TGGCTTCTAGCGCGAAGAACACAGCTGGGCTCTAGGGG	101				

```

Db      248 TGGCTTCTGAGCGGAAGAACACCTGGGCTCTAGGGGG 288

RESULT 2
LOCUS   CF315931
DEFINITION HD-05-A13.b1 OshDAC1-overexpressing transgenic rice plasmid cDNA
          library (HD) Oryza sativa (japonica cultivar-group) cDNA clone
          HD-05-A13, mRNA sequence.
ACCESSION CF315931
VERSION   CF315931.1 GI:33687692
KEYWORDS EST.
SOURCE   Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
          Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 378)
AUTHORS   Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,
          Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
          Large-scale Sequencing Analysis of Rice ESTs
          Unpublished (2003)
JOURNAL   Contact: Nahm B.H.
          Genomics and Genetics Institute, GreenGene Biotech Inc.; Division
          of Bioscience and Bioinformatics, Myongji University
          Yongin, Kyeonggi, Korea
          Tel: 82 31 330 6193
          Fax: 82 31 321 6355
          Email: bhnahm@gbio.com, bhnahm@bio.myongji.ac.kr.
FEATURES             source
     source
     1..378
     /organism="Oryza sativa (japonica cultivar-group)"
     /mol_type="mRNA"
     /cultivar="Nackdong"
     /db_xref="taxon:39947"
     /clone="HD-05-A13"
     /tissue_type="callus"
     /dev_stage="proliferated callus on 2N6 media for 2 weeks"
     /lab_host="E.coli DH10B"
     /clone_lib="OshDAC1-overexpressing transgenic rice plasmid
     cDNA library (HD)"
     /note="Vector: PCR4-TOPO; Site 1: EcoRI; Callus was
     treated with ABA(20um) for 1hr. Oligo-capped mRNA was
     reverse transcribed and then used for PCR. mRNA was
     derived from rice Histone Deacetylase overexpression
     line."

ORIGIN
Query Match      84.2%; Score 85; DB 7; Length 378;
Best Local Similarity 100.0%; Pred. No. 6.4e-15;
Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGATCGGTGGGCTCTA 60
      |||||
Db      293 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGATCGGTGGGCTCTA 352

Qy      61 TGGCTTCTGAGCGGAAGAACACG 85
      |||||
Db      353 TGGCTTCTGAGCGGAAGAACACG 377

RESULT 3
LOCUS   CN161285
DEFINITION 950563 MARC 4P1G Sus scrofa cDNA 5', mRNA sequence.
ACCESSION CN161285
VERSION   CN161285.1 GI:46175715
KEYWORDS EST.
SOURCE   Sus scrofa (pig)
ORGANISM Sus scrofa
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 640)

AUTHORS   Smith,T.P.L., Preking,B.A., Ford,J.J., Vallet,J.L., Wise,T.A.,
          Nonneman,D.J., Wray,J.E. and Keele,J.W.
          Porcine EST collection using a normalized library constructed from
          embryos representing early developmental stages
          Unpublished (2003)
JOURNAL   Contact: Smith TPL
          USDA, ARS, US Meat Animal Research Center
          PO Box 166, Clay Center, NE 68933-0166, USA
          Tel: 402 762 4366
          Fax: 402 762 4390
          Email: smith@email.marc.usda.gov
          Single pass sequencing. Bases called with phred v0.020425.c and
          trimmed with the aid of the trim_alt option. Vector identified with
          cross_match v0.990329.
          Plate: TWM8058 row: C column: 15
          Seq primer: GTAATACGACTCACTATAGGG.
FEATURES             Location/Qualifiers
     source
     1..640
     /organism="Sus scrofa"
     /mol_type="mRNA"
     /db_xref="taxon:9823"
     /tissue_type="pooled"
     /lab_host="DH10B"
     /clone_lib="MARC 4P1G"
     /note="Vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI;
     Library made with combined RNA from day-10, day-13,
     day-15, day-25, and day-30 whole embryos."

ORIGIN
Query Match      69.9%; Score 70.6; DB 7; Length 640;
Best Local Similarity 81.2%; Pred. No. 1.5e-10;
Matches 82; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy      1 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGATCGGTGGGCTCTA 60
      |||||
Db      499 ACACCAGGGGGGGATTGGGAACAAATACCAGGCTGCGGGGAGCGGGGCTTTA 558

Qy      61 TGGCTTCTGAGCGGAAGAACACGCTGGGGCTCTAGGGGG 101
      |||||
Db      559 TGGCTTCTGAGCGGAAGAACACCTGGGGCTTTGGGGGG 599

RESULT 4
LOCUS   BX993205
DEFINITION Reverse strand read from insert in 3'HPRT insertion targeting and
          chromosome engineering clone MHP272m10, genomic survey sequence.
ACCESSION BX993205
VERSION   BX993205.1 GI:49724663
KEYWORDS GSS; genome survey sequence; MICER.
SOURCE   Mus musculus (house mouse)
ORGANISM Mus musculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
          1 (bases 1 to 333)
          Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
          Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
          Rogers,J. and Bradley,A.
          Direct Submission
          Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
          CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES             Location/Qualifiers
     source
     1..333
     /organism="Mus musculus"
     /mol_type="genomic DNA"
     /db_xref="taxon:10090"
     /clone="MHP272m10"
     /clone_lib="WHPP"

ORIGIN
Query Match      65.7%; Score 66.4; DB 9; Length 333;
Best Local Similarity 87.5%; Pred. No. 2.4e-09;
Matches 84; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

```



```

QY 1 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGATCGGTGGGCTCTA 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 109 ACAGCAAGGGGAGGAGTGGGAAGACAATAGCAGGCATGCTGGGGAAGA-GAGGGCTCTA 51
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 61 TGGCTTCTGAGCGGGAAGAACACAGCTGGGCTCTA 96
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 50 TGGCTTCTGAGCGGGAAGAACACATCTAGGCTGTGA 15
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 5
BZ851581
LOCUS BZ851581 772 bp DNA linear GSS 18-MAR-2003
DEFINITION CH240_281G22.TJ CHORI-240 Bos taurus genomic clone CH240_281G22,
genomic survey sequence.
ACCESSION BZ851581
VERSION BZ851581.1 GI:29078986
KEYWORDS GSS
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovinae; Bos.
REFERENCE 1 (bases 1 to 772)
AUTHORS Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K.,
Shwartsbeyn,A., Gebregeorgis,E., Chen,D., Riggs,F., de Jong,P.,
Crawford,A.M. and McEwan,J.C.
TITLE Bovine BAC End Sequences from Library CHORI-240
JOURNAL Unpublished (2003)
COMMENT Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@igr.org
Clones are derived from the bovine BAC library CHORI-240
(http://www.chori.org/bacpac/bovine240.htm). For BAC library
availability, please contact Pieter de Jong (pjejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/orderinginformation.htm). This work
was undertaken as part of the International Bovine BAC Mapping
Consortium (IBBMC) by AgResearch Ltd., New Zealand and The
Institute of Genomic Research (TIGR), USA.
Plate: 281 row: G column: 22
Seq primer: SP6
Class: BAC ends.
FEATURES
    source
        Location/Qualifiers
            1..772
                /organism="Bos taurus"
                /mol_type="genomic DNA"
                /strain="breed: Hereford"
                /db_xref="taxon:9913"
                /clone="CH240_281G22"
                /sex="Male"
                /cell_type="Blood"
                /clone_lib="CHORI-240"
                /note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
                Hereford bull LI Domino 99375; CHORI-240 Bovine BAC
                library (Male) produced by Pieter de Jong"

ORIGIN
    Query Match 65.3%; Score 66; DB 8; Length 772;
    Best Local Similarity 93.2%; Pred. No. 3.6e-09;
    Matches 69; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGATCGGTGGGCTCTA 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 219 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGATCGGTGGGCTCTA 278
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 61 TGGCTTCTGAGCGG 74
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 279 TGGGTACCCAGGTG 292
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 6
CK462023
LOCUS CK462023 755 bp mRNA linear EST 14-JAN-2004
DEFINITION 932791 MARC 4PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION CK462023
VERSION CK462023.1 GI:40833304
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 755)
AUTHORS Smith,T.P.L., Fekking,B.A., Ford,J.J., Vallet,J.L., Wise,T.A.,
Nonneman,D.J., Wray,J.E. and Keele,J.W.
TITLE Porcine EST collection using a normalized library constructed from
embryos representing early developmental stages
JOURNAL Unpublished (2003)
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross_match v0.990329.
Plate: TW8036 row: C column: 13
Seq primer: GTAATAGCTCACTATAGGG.
FEATURES
    source
        Location/Qualifiers
            1..755
                /organism="Sus scrofa"
                /mol_type="mRNA"
                /db_xref="taxon:9823"
                /tissue_type="pooled"
                /lab_host="DH10B"
                /clone_lib="MARC 4PIG"
                /note="Vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI;
                Library made with combined RNA from day-10, day-13,
                day-15, day-25, and day-30 whole embryos."

ORIGIN
    Query Match 62.0%; Score 62.6; DB 7; Length 755;
    Best Local Similarity 76.2%; Pred. No. 3.8e-08;
    Matches 77; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 1 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGATCGGTGGGCTCTA 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 642 ACACAGCGGGGAGGATCGGGAAGAAAACAATACCGGCTGCGCGGGAGCGGGGTTTTT 701
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 61 TGGCTTCTGAGCGGGAAGACACAGCTGGGCTCTAGGGG 101
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 702 TGGTTTTTGAAGGGGAAAAAACCCGTTGGGGCTTTAGGGGG 742
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 7
CF366325
LOCUS CF366325 683 bp mRNA linear EST 25-AUG-2003
DEFINITION 840970 MARC 3PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION CF366325
VERSION CF366325.1 GI:34169801
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 683)
AUTHORS Smith,T.P.L., Fekking,B.A., Ford,J.J., Vallet,J.L., Fox,J.,
Wise,T.A., Nonneman,D.J., Wray,J.E. and Keele,J.W.
TITLE A second set of porcine ESTs from a pooled-tissue normalized
library
JOURNAL Unpublished (2003)

```

COMMENT Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smith@email.marc.usda.gov
 Single pass sequencing. Bases called with phred v0.020425.c and
 trimmed with the aid of the trim_alt option. Vector identified with
 cross_match v0.990329.
 Plate: SRG8023 row: B column: 14
 Seq primer: GTAATACGACTCACTATAGGG.
 Location/Qualifiers
 1..683
 /organism="Sus scrofa"
 /mol_type="mRNA"
 /db_xref="taxon:9823"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /clone_lib="MARC 3P1G"
 /note="Vector: pcDNA3.1; Site 1: EcoRI; Site 2: NotI;
 Library made with RNA pooled from multiple tissues
 including brain, liver, muscle, placenta/endometrium,
 ovary, testes, and bone marrow."
 ORIGIN
 Query Match 58.8%; Score 59.4; DB 7; Length 683;
 Best Local Similarity 74.3%; Pred. No. 3.4e-07;
 Matches 75; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
 QY 1 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGCATGCTGGGGATCGGTGGGCTCTA 60
 DB 526 ACACAGGGGGAGGTTGGAAAAAATAACAGCTTCCTGGGGATCCGGGGGCTTT 585
 QY 61 TGGCTTCTGAGCGGAAGAACCAGCTGGGGCTCTAGGGGG 101
 DB 586 TGGTTCTGGGGGAAAAAACCCCTCGGGGCTTTGGGGG 626
 RESULT 8
 LOCUS CB419493 648 bp mRNA linear EST 25-MAR-2003
 DEFINITION 592356 MARC 6BOV Bos taurus cDNA 5', mRNA sequence.
 ACCESSION CB419493
 VERSION CB419493.1 GI:29184608
 KEYWORDS EST.
 SOURCE Bos taurus (cow)
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovinae; Bos.
 REFERENCE 1 (bases 1 to 648)
 AUTHORS Smith,T.P.L., Roberts,A.J., Echtenkamp,S.E., Chitko-McKown,C.G.,
 Wray,J.E. and Keele,J.W.
 TITLE A second set of bovine ESTs from pooled-tissue normalized libraries
 JOURNAL Unpublished (2003)
 COMMENT Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smith@email.marc.usda.gov
 Single pass sequencing. Bases called with phred v0.020425.c and
 trimmed with the aid of the trim_alt option. Vector identified with
 cross_match v0.990329.
 Plate: FOY8013 row: P column: 9
 Seq primer: GTAATACGACTCACTATAGGG.
 Location/Qualifiers
 1..648
 /organism="Bos taurus"
 /mol_type="mRNA"
 /db_xref="taxon:9913"
 /tissue_type="pooled"
 /lab_host="DH10B"

/clone_lib="MARC 6BOV"
 /note="Vector: pcDNA3.1; Site 1: EcoRI; Site 2: NotI;
 Library made with RNA pooled from multiple tissues
 including liver, lung, hypothalamus, pituitary, and
 placenta/endometrium."
 ORIGIN
 Query Match 58.0%; Score 58.6; DB 6; Length 648;
 Best Local Similarity 75.3%; Pred. No. 5.8e-07;
 Matches 73; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
 QY 5 CAAGGGGAGGATTGGGAAGACAATAGCAGCATGCTGGGGATCGGTGGGCTCTATGGC 64
 DB 539 CCAGGGGGAGGTTTGGAAAACTTACCGGCTCTCTGGGGATCCGGGGGCTTTTGGT 598
 QY 65 TTCTAGGGCGAAAGAACCCAGCTGGGGCTCTAGGGGG 101
 DB 599 TTTTGGCGGAAAAAACCACTTGGGGCTTTAGGGGG 635
 RESULT 9
 LOCUS CR010518/c 666 bp DNA linear GSS 05-JUL-2004
 DEFINITION Forward strand read from insert in 3'HPRT insertion targeting and
 chromosome engineering clone MHP232a16, genomic survey sequence.
 ACCESSION CR010518
 VERSION CR010518.1 GI:49743509
 KEYWORDS GSS; genome survey sequence; MICER.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 666)
 AUTHORS Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,I.,
 Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
 Rogers,J. and Bradley,A.
 TITLE Direct Submission
 JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. http://www.sanger.ac.uk/MICER
 FEATURES
 source
 1..666
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /clone="MHP232a16"
 /clone_lib="MHPp"
 ORIGIN
 Query Match 53.9%; Score 54.4; DB 9; Length 666;
 Best Local Similarity 72.9%; Pred. No. 1.1e-05;
 Matches 70; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
 QY 2 CAGCAAGGGGAGGATTGGGAAGACAATAGCAGCATGCTGGGGATCGGTGGGCTCTAT 61
 DB 110 CAGCAGGGGGAGGATTAAGAGGACATATACAGCATGTTGGCAGGACGGTGGGTTTC 51
 QY 62 GGCTTCTGAGCGGAAAGAACCCAGCTGGGGCTCTAG 97
 DB 50 CGAGTATGCGTGGAAAGATCCAGCAGGGGCTGGAG 15
 RESULT 10
 LOCUS CR046705/c 754 bp DNA linear GSS 05-JUL-2004
 DEFINITION Forward strand read from insert in 3'HPRT insertion targeting and
 chromosome engineering clone MHP79m05, genomic survey sequence.
 ACCESSION CR046705
 VERSION CR046705.1 GI:49779760
 KEYWORDS GSS; genome survey sequence; MICER.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 754)
 AUTHORS Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L., Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y., Rogers,J. and Bradley,A.
 TITLE Direct Submission
 JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. <http://www.sanger.ac.uk/MICER>
 FEATURES
 source
 Location/Qualifiers
 1..754
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /clone="MHPP79m05"
 /clone_lib="MHPP"

ORIGIN
 Query Match 48.7%; Score 49.2; DB 9; Length 754;
 Best Local Similarity 70.2%; Pred. No. 0.00039;
 Matches 66; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
 3 AGCAAGGGGGAGGATTGGGAAGACAAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATG 62
 |||||
 108 AGCAGGGGAAGTATGCGCAGACATAGCTGTCCCTCCGCGTTGGGAAGGATCTTTT 49
 |||||
 63 GCTTCTGAGCGGGAAGAACACAGCTGGGGCTCTA 96
 |||||
 48 GCTTTCGTGGCGGAAGACCCAGGTGCGGATTTA 15
 |||||

RESULT 11
 CR159589/c
 LOCUS
 DEFINITION Forward strand read from insert in 3'HPRT insertion targeting and chromosome engineering clone MHP67p12, genomic survey sequence.
 ACCESSION CR159589
 VERSION CR159589.1 GI:49938438
 KEYWORDS GSS; genome survey sequence; MICER.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 119)
 Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L., Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y., Rogers,J. and Bradley,A.
 TITLE Direct Submission
 JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. <http://www.sanger.ac.uk/MICER>
 FEATURES
 source
 Location/Qualifiers
 1..119
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /clone="MHP67p12"
 /clone_lib="MHPP"

ORIGIN
 Query Match 47.7%; Score 48.2; DB 9; Length 119;
 Best Local Similarity 69.9%; Pred. No. 0.00056;
 Matches 65; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
 2 CAGCAAGGGGGAGGATTGGGAAGACAAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTAT 61
 |||||
 110 CTGCTTGGGGGTGGTCTGGTGTCTTCTGCTGGCTGCGGTGGGCTCTTC 51
 |||||
 62 GGCCTTCTGAGCGGGAAGAACAGCTGGGGCTC 94
 |||||
 50 GGCCTCTGGGGGTTTGTCTCTGCTGGGGCTC 18
 |||||

RESULT 12
 CB763949
 LOCUS

DEFINITION AMGNNUC:MRBE3-00121-F5-A rat brain E15 (10374) Rattus norvegicus cDNA clone mrbe3-00121-f5 5', mRNA sequence.
 ACCESSION CB763949
 VERSION CB763949.1 GI:29852340
 KEYWORDS EST.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 1 (bases 1 to 422)
 Amgen EST Program.
 Amgen Rat EST Program
 Unpublished (2003)
 Contact: Dan Fitzpatrick
 Amgen, Inc
 One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
 Tel: 805 447-4881
 Plate: 00121 row: f column: 5.
 Location/Qualifiers
 1..422
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /clone="mrbe3-00121-f5"
 /tissue_type="brain E15"
 /clone_lib="rat brain E15 (10374)"
 /note="Vector: pBCB; Site_1: BatXI; Site_2: NotI; rat brain E15"

ORIGIN
 Query Match 47.7%; Score 48.2; DB 6; Length 422;
 Best Local Similarity 67.3%; Pred. No. 0.00071;
 Matches 68; Conservative 0; Mismatches 33; Indels 0; Gaps 0;
 1 ACAGCAAGGGGGAGGATTGGGAAGACAAATAGCAGGCATGCTGGGGATGCGGTGGGCTCTA 60
 |||||
 286 ATAGTAAGGGGGAGGACCGCGGAAGATACAGTAGTACGTCGGGAGCGTGGCGGCTCTCA 345
 |||||
 61 TGGCTTCTGAGCGGGAAGACACAGCTGGGGCTCTAGGGGG 101
 |||||
 346 CGGTCTCGAGGTGGAAGAATTAGTCGGGGCTCTGAGGGGG 386
 |||||

RESULT 13
 CB760260
 LOCUS
 DEFINITION AMGNNUC:MRBE3-00121-G12-A rat brain E15 (10374) Rattus norvegicus cDNA clone mrbe3-00121-g12 5', mRNA sequence.
 ACCESSION CB760260
 VERSION CB760260.1 GI:29848651
 KEYWORDS EST.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 1 (bases 1 to 433)
 Amgen EST Program.
 Amgen Rat EST Program
 Unpublished (2003)
 Contact: Dan Fitzpatrick
 Amgen, Inc
 One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
 Tel: 805 447-4881
 Plate: 00121 row: g column: 12.
 Location/Qualifiers
 1..433
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /clone="mrbe3-00121-g12"
 /tissue_type="brain E15"

```

/clone.lib="rat brain E15 (10374)"
/note="Vector: pBCB; Site_1: BstXI; Site_2: NotI; rat
brain E15"

ORIGIN
Query Match          47.7%; Score 48.2; DB 6; Length 433;
Best Local Similarity 67.3%; Pred. No. 0.00071;
Matches 68; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy 1 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATCGGTGGGCTCTCA 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 286 ATAGTAAGGGGAGGACCGCGGAAGATAACAGTAGGTACGTGCGGACGTCGGCGGTCTCA 345

Qy 61 TGGCTTCTGAGCGGAAGAACACCAGCTGGGGCTCTAGGGGG 101
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 346 CGGTCTCGAGGTGGAAAGATTAGTCGGGGTCTGAGGGGG 386

RESULT 14
CB759281
LOCUS
DEFINITION
  AMGNNUC:MRBE3-00120-C9-A rat brain E15 (10374) Rattus norvegicus
  cDNA clone mrbe3-00120-c9 5', mRNA sequence.
ACCESSION
  CB759281
VERSION
  CB759281.1 GI:29847672
KEYWORDS
  EST.
SOURCE
  Rattus norvegicus (Norway rat)
ORGANISM
  Rattus norvegicus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
  Rattus.
REFERENCE
  1 (bases 1 to 434)
  Angen EST Program.
  TITLE
    Angen Rat EST Program
  JOURNAL
    Unpublished (2003)
  COMMENT
    Contact: Dan Fitzpatrick
    Angen, Inc
    One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA
    Tel: 805 447-4881
    Plate: 00120 row: c column: 9.
  Location/Qualifiers
    1..434
    /organism="Rattus norvegicus"
    /mol_type="mRNA"
    /db_xref="taxon:10116"
    /clone="mrbe3-00120-C9"
    /tissue_type="brain E15"
    /clone_lib="rat brain E15 (10374)"
    /note="Vector: pBCB; Site_1: BstXI; Site_2: NotI; rat
    brain E15"

ORIGIN
Query Match          47.7%; Score 48.2; DB 6; Length 434;
Best Local Similarity 67.3%; Pred. No. 0.00071;
Matches 68; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy 1 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATCGGTGGGCTCTCA 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 274 ATAGTAAGGGGAGGACCGCGGAAGATAACAGTAGGTACGTGCGGACGTCGGCGGTCTCA 333

Qy 61 TGGCTTCTGAGCGGAAGAACACCAGCTGGGGCTCTAGGGGG 101
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 334 CGGTCTCGAGGTGGAAAGATTAGTCGGGGTCTGAGGGGG 374

RESULT 15
CB742761
LOCUS
DEFINITION
  AMGNNUC:MRBE3-00121-C7-A rat brain E15 (10374) Rattus norvegicus
  cDNA clone mrbe3-00121-c7 5', mRNA sequence.
ACCESSION
  CB742761
VERSION
  CB742761.1 GI:29810059
KEYWORDS
  EST.
```

```

Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
  1 (bases 1 to 449)
  Angen EST Program.
  AUTHORS
    Angen Rat EST Program
  JOURNAL
    Unpublished (2003)
  COMMENT
    Contact: Dan Fitzpatrick
    Angen, Inc
    One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA
    Tel: 805 447-4881
    Plate: 00121 row: c column: 7.
  Location/Qualifiers
    1..449
    /organism="Rattus norvegicus"
    /mol_type="mRNA"
    /db_xref="taxon:10116"
    /clone="mrbe3-00121-C7"
    /tissue_type="brain E15"
    /clone_lib="rat brain E15 (10374)"
    /note="Vector: pBCB; Site_1: BstXI; Site_2: NotI; rat
    brain E15"

ORIGIN
Query Match          47.7%; Score 48.2; DB 6; Length 449;
Best Local Similarity 67.3%; Pred. No. 0.00071;
Matches 68; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy 1 ACAGCAAGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATCGGTGGGCTCTCA 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 289 ATAGTAAGGGGAGGACCGCGGAAGATAACAGTAGGTACGTGCGGACGTCGGCGGTCTCA 348

Qy 61 TGGCTTCTGAGCGGAAGAACACCAGCTGGGGCTCTAGGGGG 101
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 349 CGGTCTCGAGGTGGAAAGATTAGTCGGGGTCTGAGGGGG 389

Search completed: July 14, 2005, 23:22:58
Job time : 968.667 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:39:07 ; Search time 756.618 Seconds
(without alignments)
6468.225 Million cell updates/sec

Title: US-09-482-682-43_COPY_8283_8383
Perfect score: 101
Sequence: 1 aggggtattgtctcatgagc.....Gaaagtgcacctgagctc 101

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_scs.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	101	100.0	142	6 AR356490	AR356490 Sequence
C 2	101	100.0	142	6 AR538046	AR538046 Sequence
C 3	101	100.0	228	6 E00019	E00019 DNA coding
C 4	101	100.0	240	1 PMOENDO	M10199 Plasmid pMm
C 5	101	100.0	251	6 E00018	E00018 DNA coding
C 6	101	100.0	251	6 I01644	I01644 Sequence 1
C 7	101	100.0	344	11 HUMUT5345	LI8624 Human chrom
C 8	101	100.0	400	6 BD195256	BD195256 Nucleotid
C 9	101	100.0	456	6 E00892	E00892 Synthetic D
C 10	101	100.0	456	6 E01156	E01156 DNA fragmen
C 11	101	100.0	456	6 E01274	E01274 DNA encodin
C 12	101	100.0	456	6 E01302	E01302 DNA encodin
C 13	101	100.0	466	6 AX260098	AX260098 Sequence
C 14	101	100.0	573	6 AX260150	AX260150 Sequence
C 15	101	100.0	693	6 A43586	A43586 Sequence 11
C 16	101	100.0	693	6 AR116755	AR116755 Sequence
C 17	101	100.0	998	1 AY559171	AY559171 Pseudomon
C 18	101	100.0	1011	1 SMTMAQGE	X97254 S.marcescen
C 19	101	100.0	1012	2 CEC11F10	Z92776 Caenorhabdi

20	101	100.0	1014	4 CFAJ4121	AJ224121 Canis fam
C 21	101	100.0	1027	1 AY589493	AY589493 Escherich
C 22	101	100.0	1040	1 AY538698	AY538698 Serratia
C 23	101	100.0	1040	1 AY538700	AY538700 Serratia
C 24	101	100.0	1040	1 AY538701	AY538701 Serratia
C 25	101	100.0	1040	1 AY538702	AY538702 Serratia
C 26	101	100.0	1041	1 AY538699	AY538699 Serratia
C 27	101	100.0	1042	1 AY394610	AJ394610 Klebsiell
C 28	101	100.0	1042	1 ECO308558	AJ308558 Escherich
C 29	101	100.0	1044	1 AY392531	AJ392531 Streptoco
C 30	101	100.0	1044	1 AY452662	AY452662 Streptoco
C 31	101	100.0	1054	1 AF104441	AF104441 Klebsiell
C 32	101	100.0	1054	1 AF104442	AF104442 Escherich
C 33	101	100.0	1058	6 I03356	I03356 Sequence 4
C 34	101	100.0	1064	1 AY628199	AY628199 Escherich
C 35	101	100.0	1069	1 AF535127	AF535127 Klebsiell
C 36	101	100.0	1069	1 AY243512	AY243512 Klebsiell
C 37	101	100.0	1071	1 AY628175	AY628175 Escherich
C 38	101	100.0	1072	1 AY101764	AY101764 Klebsiell
C 39	101	100.0	1073	6 AR371489	AR371489 Sequence
C 40	101	100.0	1073	6 AX195443	AX195443 Sequence
C 41	101	100.0	1075	1 AY729027	AY729027 Proteus m
C 42	101	100.0	1075	1 PATN1PN2	X54606 Pseudomonas
C 43	101	100.0	1075	1 PATN2PN1B	X54607 Pseudomonas
C 44	101	100.0	1075	1 PATN3PN1A	X54604 Pseudomonas
C 45	101	100.0	1080	1 AF027199	AF027199 Klebsiell

ALIGNMENTS

RESULT 1
AR356490/c
LOCUS AR356490 142 bp DNA linear PAT 17-AUG-2003
DEFINITION Sequence 2608 from patent US 6593114.
ACCESSION AR356490
VERSION AR356490.1 GI:33762574
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 142)
AUTHORS Kunsch,C.A., Choi,G.H., Barash,S., Dillon,P.J., Fannon,M.R. and Rosen,C.A.
TITLE Staphylococcus aureus polynucleotides and sequences
JOURNAL Patent: US 6593114-A 2608 15-JUL-2003;
FEATURES Location/Qualifiers
source 1..142
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 142;
Best Local Similarity 100.0%; Pred. No. 8,7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGATGTATTAGAAATAAACAAATAG 60
Db 107 AGGGTTATTGCTCATGAGCGGATACATATTGATGTATTAGAAATAAACAAATAG 48
Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 101
Db 47 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 7
RESULT 2
AR538046/c
LOCUS AR538046 142 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 2608 from patent US 6737248.
ACCESSION AR538046
VERSION AR538046.1 GI:53929263
KEYWORDS
SOURCE Unknown.

```
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 142)
AUTHORS Kunsch,C.A., Choi,G.A., Barash,S.C., Dillon,P.J., Fannon,M.R. and
Rosen,C.A.
TITLE Staphylococcus aureus polynucleotides and sequences
JOURNAL Patent: US 6737248-A 2408 18-MAY-2004;
FEATURES Location/Qualifiers
source 1..142
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 142;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGAATGTATTAGAAAAATAACAAATAG 60
Db 107 AGGGTTATTGTCATGAGCGGATACATATTGAATGTATTAGAAAAATAACAAATAG 48
Qy 61 GGGTTCGCGGCACATTTCCCGAAAAAGTGCCACCTGACGTC 101
Db 47 GGGTTCGCGGCACATTTCCCGAAAAAGTGCCACCTGACGTC 7
RESULT 3
E00019/c
LOCUS DNA coding for Escherichia coli penicillinase.
ACCESSION E00019
VERSION E00019.1 GI:2168327
KEYWORDS JP 1981154999-A/2.
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
1 (bases 1 to 228)
Uorutaa,G. and Karen,T.
SYNTHESIS OF SELECTIVE AGED PROTEIN OR POLYPEPTIDE IN BACTERIA
JOURNAL Patent: JP 1981154999-A 2 30-NOV-1981;
UNIV HARVARD
OS Escherichia coli
PN JP 1981154999-A/2
PD 30-NOV-1981
PF 09-APR-1981 JP 1981052488
PR 11-APR-1980 US 80 139225
PI UORUTAA GIRUBAATO, KAREN TARUMATSUJI
PC C12P21/00,C07H21/00,C12N1/00,C12N15/00//C12R1/19; CC
strandedness: Double;
CC topology: Linear;
CC anti-sense: No;
CC *source: clone=pKT218;
FH Key Location/Qualifiers
FH CDS 210..>228
FT /product='E.coli penicillinase'.
FEATURES
source 1..228
Location/Qualifiers
/organism="Escherichia coli"
/mol_type="genomic DNA"
/db_xref="taxon:562"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 228;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGAATGTATTAGAAAAATAACAAATAG 60
Db 175 AGGGTTATTGTCATGAGCGGATACATATTGAATGTATTAGAAAAATAACAAATAG 116
Qy 61 GGGTTCGCGGCACATTTCCCGAAAAAGTGCCACCTGACGTC 101
```

```
Db 115 GGGTTCGCGGCACATTTCCCGAAAAAGTGCCACCTGACGTC 75
RESULT 4
PMOENDO/c
LOCUS Plasmid pMM110 region of endo VII cleavage sites near cruciform
DEFINITION structures.
ACCESSION M10199
VERSION M10199.1 GI:150826
KEYWORDS
SOURCE Plasmid pMM110
ORGANISM Plasmid pMM110
other sequences; plasmids.
REFERENCE 1 (bases 1 to 240)
AUTHORS Kemper,B., Jensch,F., von Depka-Prondzynski,M., Fritz,H.J.,
Borgmeyer,U. and Mizuuchi,K.
TITLE Resolution of Holliday structures by endonuclease VII as observed
in interactions with cruciform DNA
JOURNAL Cold Spring Harb. Symp. Quant. Biol. 49, 815-825 (1984)
MEDLINE 85153063
PUBMED 6397324
COMMENT Original source text: Plasmid pMM110 DNA.
FEATURES Location/Qualifiers
source 1..240
/organism="Plasmid pMM110"
/mol_type="genomic DNA"
/db_xref="taxon:2599"
/plasmid="Plasmid pMM110"
ORIGIN Unreported.
Query Match 100.0%; Score 101; DB 1; Length 240;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGAATGTATTAGAAAAATAACAAATAG 60
Db 151 AGGGTTATTGTCATGAGCGGATACATATTGAATGTATTAGAAAAATAACAAATAG 92
Qy 61 GGGTTCGCGGCACATTTCCCGAAAAAGTGCCACCTGACGTC 101
Db 91 GGGTTCGCGGCACATTTCCCGAAAAAGTGCCACCTGACGTC 51
RESULT 5
E00018/c
LOCUS DNA coding for Escherichia coli penicillinase.
ACCESSION E00018
VERSION E00018.1 GI:2168326
KEYWORDS JP 1981154999-A/1.
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
1 (bases 1 to 251)
Uorutaa,G. and Karen,T.
SYNTHESIS OF SELECTIVE AGED PROTEIN OR POLYPEPTIDE IN BACTERIA
JOURNAL Patent: JP 1981154999-A 1 30-NOV-1981;
UNIV HARVARD
OS Escherichia coli
PN JP 1981154999-A/1
PD 30-NOV-1981
PF 09-APR-1981 JP 1981052488
PR 11-APR-1980 US 80 139225
PI UORUTAA GIRUBAATO, KAREN TARUMATSUJI
PC C12P21/00,C07H21/00,C12N1/00,C12N15/00//C12R1/19; CC
strandedness: Double;
CC topology: Linear;
CC anti-sense: No;
CC fragment type: N-Terminal Fragment;
CC *source: clone=pKT241;
```

EH Key Location/Qualifiers
FH CDS 210..>252
FT /product='E.coli penicillinase' FT
TATA_signal 190..196.
Location/Qualifiers
1..251
/organism='Escherichia coli'
/mol_type='genomic DNA'
/db_xref='taxon:562'

FEATURES
source
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 251;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 60
|||||
DB 175 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 116
|||||

QY 61 GGGTTCCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 101
|||||
DB 115 GGGTTCCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 75
|||||

RESULT 6
I01644/c
LOCUS
DEFINITION Sequence 1 from Patent US 4338397.
ACCESSION I01644
VERSION I01644.1 GI:267685
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 251)
AUTHORS Gilbert, W. and Talmadge, K.
TITLE Mature protein synthesis
JOURNAL Patent: US 4338397-A 1 06-JUL-1982;
President and Fellows of Harvard College; Cambridge, MA

FEATURES
source
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 251;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 60
|||||
DB 175 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 116
|||||

QY 61 GGGTTCCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 101
|||||
DB 115 GGGTTCCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 75
|||||

RESULT 7
HUMUT5345
LOCUS
DEFINITION Human chromosome 8 STS UT5345, sequence tagged site.
ACCESSION L18624
VERSION L18624.1 GI:308338
KEYWORDS STS; PCR primer; STS sequence; microsatellite marker; microsatellite repeat; repeat polymorphism; sequence tagged site.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 344)
AUTHORS Gerken, S.C., Matsunami, N., Lawrence, E., Carlson, M., Moore, M.,

Ballard, L., Melis, R., Robertson, M., Bradley, P., Elsner, T., Tingey, A., Rodriguez, P., Albertsen, H., Lalouel, J.-M. and White, R. Genetic and physical mapping of simple sequence repeat containing sequence tagged sites from the human genome
Unpublished (1993)
Original source text: Homo sapiens DNA.
Submitted by: Utah Center for Human Genome Research University of Utah, Dept. of Human Genetics
2160 Eccles Institute of Human Genetics
Salt Lake City, UT 84112
e-mail: sts@corona.med.utah.edu
Primer A: GAGCAAAAACAGCAGGCAAAATGC
Primer B: TTCGGGAATGTGCGGGAAC
32P-label: B Primer
PCR Profile:
Initial Denaturation: 94C 300sec
PCR Cycles: 30
Denaturation: 94C 10sec
Annealing: 60C 10sec
Extension: 72C 20sec
Mg++: 2mM
Gel: Acrylamide 7%, Formamide 32%, Urea 34%
Alleles: 2

FEATURES
source
1..344
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'
/map='8'
36..224
/standard_name='STS UT5345'
36..60
primer_bind
primer_bind complement (202..224)

ORIGIN
Query Match 100.0%; Score 101; DB 11; Length 344;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 60
|||||
DB 141 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 200
|||||

QY 61 GGGTTCCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 101
|||||
DB 201 GGGTTCCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 241
|||||

RESULT 8
BD195256/c
LOCUS
DEFINITION Nucleotide sequence of Escherichia coli pathogenicity islands.
ACCESSION BD195256
VERSION BD195256.1 GI:33005021
KEYWORDS JP 2002513277-A/43.
SOURCE unidentified
ORGANISM unidentified
REFERENCE
1 (bases 1 to 400)
AUTHORS Dillon, P.J., Choi, G.H. and Welch, R.A.
TITLE Nucleotide sequence of Escherichia coli pathogenicity islands
JOURNAL Patent: JP 2002513277-A 43 08-MAY-2002;
HUMAN GENOME SCIENCES INC, WISCONSIN ALUMNI RESEARCH FOUNDATION

COMMENT
OS Unidentified
PN JP 2002513277-A/43
PD 08-MAY-2002
PF 21-NOV-1997 JP 1998523916
PR 22-NOV-1996 US 60/031626.14-OCT-1997 US 60/061953 PI
PATRICK J DILLON, GIL H CHOI, RODNEY A WELCH
PC C12N15/11, C12N15/63, C07K16/12, G01N33/569, G06F17/30, G11B7/00 CC
Strandedness: Double;
CC Topology: Linear;
CC Nucleotide sequence of Escherichia coli pathogenicity islands

```

FH Key Location/Qualifiers
FT source 1..400
FT Location/Qualifiers
FEATURES
source 1..400
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 400;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 60
Db 165 AGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 106

Qy 61 GGGTCCGCGCACATTTCCCGAAAGTGCCACCTGAGTC 101
Db 105 GGGTCCGCGCACATTTCCCGAAAGTGCCACCTGAGTC 65

RESULT 9
E00892/c
LOCUS E00892 456 bp DNA linear PAT 29-SEP-1997
DEFINITION Synthetic DNA encoding fused polypeptide between E coli
beta-lactamase and human beta-urogastrone.
ACCESSION E00892
VERSION E00892.1 GI:2169153
KEYWORDS JP 1986149089-A/1.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 456)
AUTHORS Okai, H., Momota, Y., Kumakura, T., Tochifusa, N., Kitazawa, T.,
Ojida, K. and Matsushiro, A.
TITLE POLYPEPTIDE SECRETION DEVELOPMENT VECTOR, MICROORGANISM TRANSFORMED
WITH SAME, AND PRODUCTION OF POLYPEPTIDE WITH MICROORGANISM
JOURNAL Patent: JP 1986149089-A 1 07-JUL-1986;
EARTH CHEM CORP LTD
COMMENT OS Artificial gene
OC Artificial sequence; Genes.
PN JP 1986149089-A/1
PD 07-JUL-1986
PF 21-DEC-1984 JP 1984271206
PI OKAI HIDEO, MOMOTA YUTAKA, KUMAKURA TAKESHI,
TOCHIFUSA NORIYUKI,
KITAZAWA TOSHIKI, OJIDA KAZUHIRO, MATSUSHIRO AIZO
PC C12N15/00, C12N1/20, C12P21/00, (C12N1/20, C12R1:19), (C12P21/00, PC
C12R1:19);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: strain=HB101;
CC *source: clone=pVG201;
CC Feature is identified by experimental;
FH key Location/Qualifiers
FH promoter 125..170
FT of beta-lactamase
FT RBS 200..203
FT CDS 209..438
FT /product='beta-urogastrone precursor' FT
FT sig_peptide 209..277
FT /product='signal peptide of beta-lactonase' FT
FT mat_peptide 278..435
FT /product='beta-urogastrone mature peptide'.
FT Location/Qualifiers
FEATURES
source 1..456
/organism="synthetic construct"
/mol_type="genomic DNA"

FH Key Location/Qualifiers
FT source 1..400
FT Location/Qualifiers
FEATURES
source 1..400
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 60
Db 173 AGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 114

Qy 61 GGGTCCGCGCACATTTCCCGAAAGTGCCACCTGAGTC 101
Db 113 GGGTCCGCGCACATTTCCCGAAAGTGCCACCTGAGTC 73

RESULT 10
E01156/c
LOCUS E01156 456 bp DNA linear PAT 29-SEP-1997
DEFINITION DNA fragment which secrets beta urogastrone.
ACCESSION E01156
VERSION E01156.1 GI:2169415
KEYWORDS JP 1987083890-A/1.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 456)
AUTHORS Yoshikawa, K., Momota, Y., Kajifusa, N., Koide, T. and Okai, H.
TITLE POLYPEPTIDE SECRETION EXPRESSION VECTOR, MICROORGANISM TRANSFORMED
BY SAID VECTOR AND PRODUCTION OF POLYPEPTIDE USING SAID
JOURNAL Patent: JP 1987083890-A 1 17-APR-1987;
EARTH CHEM CORP LTD
COMMENT OS Artificial gene
OC Artificial sequence; Genes.
PN JP 1987083890-A/1
PD 17-APR-1987
PF 09-OCT-1985 JP 1985225393
PI YOSHIKAWA KAZUTOSHI, MOMOTA YUTAKA, KAJIFUSA NORIYUKI, PI
KOIDE TAKAO,
OKAI HIDEO
PC C12N15/00, C12N1/20, C12P21/00, (C12N1/20, C12R1:19), (C12N1/20, PC
C12R1:125);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: clone=pUG201;
FH key Location/Qualifiers
FH promoter 125..170
FT /note='beta lactamase promoter' FT RBS
FT CDS 209..439
FT /product='beta urogastrone'
FT sig_peptide 209..277
FT mat_peptide 278..436
FT /product='beta urogastrone'.
FT Location/Qualifiers
FEATURES
source 1..456
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 60
Db 173 AGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 114

Qy 61 GGGTCCGCGCACATTTCCCGAAAGTGCCACCTGAGTC 101

```



```
Db 113 GGGTTCCGGCGCACATTTCCCGAAAGTGCCACCTGACGTC 73
|||||
RESULT 11
E01274/c
LOCUS E01274 456 bp DNA linear PAT 29-SEP-1997
DEFINITION DNA encoding beta-urogastron fused with DNA encoding a promoter and
signal peptide of beta-lactamase.
ACCESSION E01274
VERSION E01274.1 GI:2169533
KEYWORDS JP 1987179398-A/1
SOURCE synthetic construct
ORGANISM
REFERENCE 1 (bases 1 to 456)
AUTHORS Okai,H., Kumakura,T., Kawamoto,S., Adachi,S., Matsubara,A.,
Ojida,K., Yano,M., Mihara,S., Matsuhiro,A. and Yanaiharu,N.
TITLE PRODUCTION OF BETA-UROGASTRONE
JOURNAL Patent: JP 1987179398-A 1 06-AUG-1987;
COMMENT EARTH CHEM CORP LTD
OS Artificial gene
OC Artificial sequence; Genes.
OS Homo sapiens
PN JP 1987179398-A/1
PD 06-AUG-1987
PF 31-JAN-1986 JP 1986021032
PI OKAI HIDEO, KUMAKURA TAKESHI, KAWAMOTO SHOICHI, ADACHI SHOJI,
MATSUBARA AKIMASA, OJIDA KAZUHIDE, YANO MAKI, MIHARA SHIGERU,
MATSUSHIRO AIZO, YANAIHARA NOBORU
PC C12P21/00,C12N15/00,(C12P21/00,C12R1:91);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key
FT Promoter 125..170
FT RBS 200..203
FT sig_peptide 209..277
FT mat_peptide 278..436
FT CDS 209..439
FT /product='beta-urogastron'
FT /product='beta-urogastron'
FEATURES
source
1..456
/organism='synthetic construct'
/mol_type='genomic DNA'
/db_xref='taxon:32630'
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTTAGAAAAATAACAAATAG 60
|||||
Db 173 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTTAGAAAAATAACAAATAG 114
|||||
Qy 61 GGGTTCCGGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101
|||||
Db 113 GGGTTCCGGCGCACATTTCCCGAAAGTGCCACCTGACGTC 73
|||||
RESULT 12
E01302/c
LOCUS E01302 456 bp DNA linear PAT 29-SEP-1997
DEFINITION DNA encoding human beta-urogastrone fused with DNA encoding
promoter and signal peptide of beta-lactamase.
ACCESSION E01302
VERSION E01302.1 GI:2169561
KEYWORDS JP 1987190083-A/1
SOURCE synthetic construct
ORGANISM
REFERENCE 1 (bases 1 to 456)
AUTHORS Deak,P., Glover,D.M. and Midgley,C.
TITLE Cell cycle progression proteins
JOURNAL Patent: WO 0172774-A 60 04-OCT-2001;
Cyclacel Limited (GB)
FEATURES
source
1..456
/organism='Drosophila melanogaster'
/mol_type='unassigned DNA'
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTTAGAAAAATAACAAATAG 60
|||||
Db 173 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTTAGAAAAATAACAAATAG 114
|||||
Qy 61 GGGTTCCGGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101
|||||
Db 113 GGGTTCCGGCGCACATTTCCCGAAAGTGCCACCTGACGTC 73
|||||
RESULT 13
AX260098/c
LOCUS AX260098 466 bp DNA linear PAT 26-OCT-2001
DEFINITION Sequence 60 from Patent WO0172774.
ACCESSION AX260098
VERSION AX260098.1 GI:16509129
KEYWORDS
SOURCE
ORGANISM Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1
AUTHORS Deak,P., Glover,D.M. and Midgley,C.
TITLE Cell cycle progression proteins
JOURNAL Patent: WO 0172774-A 60 04-OCT-2001;
Cyclacel Limited (GB)
FEATURES
source
1..466
/organism='Drosophila melanogaster'
/mol_type='unassigned DNA'
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTTAGAAAAATAACAAATAG 60
|||||
Db 173 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTTAGAAAAATAACAAATAG 114
|||||
Qy 61 GGGTTCCGGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101
|||||
Db 113 GGGTTCCGGCGCACATTTCCCGAAAGTGCCACCTGACGTC 73
|||||
```

```

/db_xref="taxon:7227"

ORIGIN
Query Match      100.0%; Score 101; DB 6; Length 466;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 60
    |||
Db 280 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 221
    |||

Qy 61 GGGTTCCGCGACATTTCCCGAAAAGTCCACCTGACGTC 101
    |||
Db 220 GGGTTCCGCGACATTTCCCGAAAAGTCCACCTGACGTC 180
    |||

RESULT 14
LOCUS AX260150 573 bp DNA linear PAT 26-OCT-2001
DEFINITION Sequence 112 from Patent WO0172774.
ACCESSION AX260150
VERSION AX260150.1 GI:16509172
KEYWORDS
SOURCE
ORGANISM Drosophila melanogaster (fruit fly)
          Drosophila melanogaster
          Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
          Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
          Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 Deak, P., Glover, D.M. and Midgley, C.
  Cell cycle progression proteins
  Patent: WO 0172774-A 112 04-OCT-2001;
  Cyclacel Limited (GB)
FEATURES
    source
    1..573
    /organism="Drosophila melanogaster"
    /mol_type="unassigned DNA"
    /db_xref="taxon:7227"

ORIGIN
Query Match      100.0%; Score 101; DB 6; Length 573;
Best Local Similarity 100.0%; Pred. No. 8.8e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 60
    |||
Db 355 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 296
    |||

Qy 61 GGGTTCCGCGACATTTCCCGAAAAGTCCACCTGACGTC 101
    |||
Db 295 GGGTTCCGCGACATTTCCCGAAAAGTCCACCTGACGTC 255
    |||

RESULT 15
LOCUS A43586 693 bp DNA linear PAT 06-MAR-1997
DEFINITION Sequence 11 from Patent WO9507357.
ACCESSION A43586
VERSION A43586.1 GI:2298779
KEYWORDS
SOURCE
ORGANISM Cuphea lanceolata
          Cuphea lanceolata
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
          rosids; Myrtales; Lythraceae; Cuphea.
REFERENCE
1 (bases 1 to 693)
  Toepfer, R., Bautor, J., Bothmann, H., Filsak, E.,
  Hoerliche-Grandpierre, C., Klein, B., Martini, N., Mueller, A.,
  Schulte, W., Voetz, M., Walek, J. and Schell, J.
  PROMOTERS
  Patent: WO 9507357-A 11 16-MAR-1995;
  MAX PLANCK GESELLSCHAFT (DE)
  Other publication CA 2169093 950316
  COMMENT
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:35:42 ; Search time 142.398 Seconds
(without alignments)
4198.742 Million cell updates/sec

Title: US-09-482-682-43_COPY_8283_8383
Perfect score: 101
Sequence: 1 aggtttattgtctatgagc.....gaaagtgcacacctgaagtc 101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	101	100.0	142	2 AAV76919	AAV76919 Staphyloc
C 2	101	100.0	228	1 AAN10032	Aan10032 Sequence
C 3	101	100.0	251	1 AAN10031	Aan10031 Sequence
C 4	101	100.0	400	2 AAV31229	AAV31229 E. coli J
C 5	101	100.0	456	1 AAN60824	Aan60824 Plasmid p
C 6	101	100.0	456	1 AAN71080	Aan71080 Sequence
C 7	101	100.0	456	1 AAN70833	Aan70833 Beta-urog
C 8	101	100.0	456	1 AAN81765	Aan81765 Sequence
C 9	101	100.0	466	6 ABA90413	ABA90413 Drosophil
C 10	101	100.0	487	2 AAX21173	Aax21173 Polynucle
C 11	101	100.0	535	2 AAX21149	Aax21149 Polynucle
C 12	101	100.0	573	6 ABA90456	ABA90456 Drosophil
C 13	101	100.0	605	12 ADH58311	Adh58311 Electroph
C 14	101	100.0	776	4 AAS30560	Aas30560 DNA encod
C 15	101	100.0	776	4 AAS27819	Aas27819 DNA encod
C 16	101	100.0	776	4 ABX42984	Abx42984 Genomic s
C 17	101	100.0	776	4 AAL07344	Aal07344 Human rep
C 18	101	100.0	776	4 AAL03229	Aal03229 Human rep
C 19	101	100.0	776	4 AAL06588	Aal06588 Human rep
C 20	101	100.0	776	4 AAL07340	Aal07340 Human rep

C 21	101	100.0	776	5 ABA14573	Abal14573 Human ner
C 22	101	100.0	776	5 AAS34681	Aas34681 Human DNA
C 23	101	100.0	776	8 ADA41574	Ada41574 Human sec
C 24	101	100.0	776	8 ACC50905	Acc50905 Human sec
C 25	101	100.0	776	8 ABZ71508	Abz71508 Secreted
C 26	101	100.0	776	9 ADB91869	Adb91869 Human sec
C 27	101	100.0	776	9 ADB61140	Adb61140 Connectiv
C 28	101	100.0	776	10 ADB94622	Adb94622 Novel hum
C 29	101	100.0	776	10 ADC74663	Adc74663 Human sec
C 30	101	100.0	776	10 ADA57709	Ada57709 BAC fragm
C 31	101	100.0	776	12 ADN41551	Adn41551 Novel hum
C 32	101	100.0	845	4 AAS30559	Aas30559 DNA encod
C 33	101	100.0	845	4 AAS27818	Aas27818 DNA encod
C 34	101	100.0	845	4 ABK42983	Abk42983 Genomic s
C 35	101	100.0	845	4 AAS41807	Aas41807 Genomic s
C 36	101	100.0	845	4 AAS41855	Aas41855 Genomic s
C 37	101	100.0	845	4 AAK85485	Aak85485 Human imm
C 38	101	100.0	845	4 AAK85434	Aak85434 Human imm
C 39	101	100.0	845	4 AAL07343	Aal07343 Human rep
C 40	101	100.0	845	4 AAL06587	Aal06587 Human rep
C 41	101	100.0	845	4 AAL07339	Aal07339 Human rep
C 42	101	100.0	845	4 AAL03228	Aal03228 Human rep
C 43	101	100.0	845	5 ABA14572	Abal14572 Human ner
C 44	101	100.0	845	5 AAS34680	Aas34680 Human DNA
C 45	101	100.0	845	9 ADB61139	Adb61139 Connectiv

ALIGNMENTS

RESULT 1

AAV76919/c
ID AAV76919 standard; DNA; 142 BP.

XX AAV76919;

DT 16-MAR-1999 (first entry)

DE Staphylococcus aureus contig SEQ ID #2608.

XX Computer readable medium; vaccine; S.aureus infection; immunodetection;
KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW skin infection; surgical wound infection; scalded skin syndrome;
KW toxic shock syndrome; ds.

XX Staphylococcus aureus.

XX EP786519-A2.

PN 30-JUL-1997.

XX 07-JAN-1997; 97EP-00100117.

XX 05-JAN-1996; 96US-0009861P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Kunsch CA, Choi GH, Barash SC, Dillon PJ, Fannon MR, Rosen CA;

XX WPI; 1997-374922/35.

XX Polynucleotide(s) and proteins derived from Staphylococcus aureus -
stored on computer readable medium and used in the production of anti-
S.aureus vaccines.

XX Claim 1; Page 2287; 3271pp; English.

XX This sequence represents one of 5191 Staphylococcus aureus DNA sequences
of the invention. The DNA sequences are recorded on a computer readable
medium, preferably selected from a floppy or hard disk, random access
memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
the S.aureus DNA sequences allows putative functions to be assigned so
that protein-encoding or regulatory regions of commercial, therapeutic or

CC industrial importance can be obtained. Specifically, sequences which are
 CC likely to encode antigens have been identified and these polypeptides can
 CC be used in a vaccine composition against *S.aureus* infection. The
 CC polypeptides can also be used in a kit for the immunodetection of
 CC *S.aureus* in a sample. *S.aureus* is implicated in numerous human diseases,
 CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,
 CC skin and surgical wound infections, scalded skin syndrome, toxic shock
 CC syndrome, etc. Organisms transformed with the DNA sequences can be used
 CC for recombinant production of the polypeptides. The new DNA sequences
 CC (and their fragments) are useful as primers or probes for isolating
 CC homologues of any of the *S.aureus* DNA sequences contained on the computer
 CC readable medium
 XX
 SQ Sequence 142 BP; 45 A; 25 C; 26 G; 45 T; 0 U; 1 Other;

Query Match 100.0%; Score 101; DB 2; Length 142;
 Best Local Similarity 100.0%; Pred. No. 2.1e-21;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 AGGGTTATTGTCATGAGCGGATACATATTGAATGATTATTAGAAAATAAACAATAG 60
 DB 107 AGGGTTATTGTCATGAGCGGATACATATTGAATGATTATTAGAAAATAAACAATAG 48
 OY 61 GGGTTCCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 101
 DB 47 GGGTTCCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 7

RESULT 2
 AAN10032/c
 ID AAN10032 standard; DNA; 228 BP.
 AC AAN10032;
 DT 13-AUG-1992 (first entry)
 DE Sequence of the pKT218 EcoRI-PstI penicillinase gene fragment.
 XX Cloning vehicle; bacterial vector; transformed host; penicillinase;
 KW insulin; ds.
 XX Escherichia coli.
 FH Key Location/Qualifiers
 FT misc_feature 1..4
 FT /*tag= a
 FT /label= sticky end
 FT misc_feature 225..228
 FT /*tag= b
 FT /label= sticky end
 XX EP38182-A.
 XX 21-OCT-1981.
 XX 09-APR-1981; 81EP-00301561.
 XX 11-APR-1980; 80US-00139225.
 XX (HARD) HARVARD COLLEGE.
 XX Gilbert W, Talmadge K;
 XX WPI; 1981-80125D/44.
 XX P-PSDB; AAP10039.
 XX Synthesis of mature protein or polypeptide - by using bacterial host
 XX transformed by cloned vehicle contg. DNA fragment etc.
 XX Example; Fig 3; 34pp; English.
 XX The closest identifiable promoter for the penicillinase gene in pKT241
 CC (AAN10031) is located in the region 14 to 20 nucleotides before its

CC translational start signal. In the examples, the 3' end of pKT241 was
 CC attached to the signal DNA sequence of the DNA fragment (19) for rat
 CC preproinsulin (see AAN10033). The closest identifiable promoter for the
 CC penicillinase gene in pKT218 (AAN10032) is located in the region 14 to 20
 CC nucleotides before its translational start signal. In the examples, the
 CC 3' end of pKT218 was attached to the signal DNA sequence of the DNA
 CC fragment (CB6) for rat preproinsulin (see AAN10034)
 XX

SQ Sequence 228 BP; 72 A; 37 C; 46 G; 73 T; 0 U; 0 Other;
 Query Match 100.0%; Score 101; DB 1; Length 228;
 Best Local Similarity 100.0%; Pred. No. 2.3e-21;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 AGGGTTATTGTCATGAGCGGATACATATTGAATGATTATTAGAAAATAAACAATAG 60
 DB 175 AGGGTTATTGTCATGAGCGGATACATATTGAATGATTATTAGAAAATAAACAATAG 116
 OY 61 GGGTTCCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 101
 DB 115 GGGTTCCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 75

RESULT 3
 AAN10031/c
 ID AAN10031 standard; DNA; 251 BP.
 XX
 AC AAN10031;
 DT 13-AUG-1992 (first entry)
 DE Sequence of the pKT241 EcoRI-PstI penicillinase gene fragment.
 XX Cloning vehicle; bacterial vector; transformed host; penicillinase;
 KW insulin; ds.
 XX Escherichia coli.
 FH Key Location/Qualifiers
 FT misc_feature 1..4
 FT /*tag= a
 FT /label= sticky end
 FT misc_feature 248..251
 FT /*tag= b
 FT /label= sticky end
 XX EP38182-A.
 XX 21-OCT-1981.
 XX 09-APR-1981; 81EP-00301561.
 XX 11-APR-1980; 80US-00139225.
 XX (HARD) HARVARD COLLEGE.
 XX Gilbert W, Talmadge K;
 XX WPI; 1981-80125D/44.
 XX P-PSDB; AAP10038.
 XX Synthesis of mature protein or polypeptide - by using bacterial host
 XX transformed by cloned vehicle contg. DNA fragment etc.
 XX Example; Fig 2; 34pp; English.
 XX The closest identifiable promoter for the penicillinase gene in pKT241
 CC (AAN10031) is located in the region 14 to 20 nucleotides before its
 CC translational start signal. In the examples, the 3' end of pKT241 was
 CC attached to the signal DNA sequence of the DNA fragment (19) for rat
 CC preproinsulin (see AAN10033). The closest identifiable promoter for the
 CC penicillinase gene in pKT218 (AAN10032) is located in the region 14 to 20
 CC nucleotides before its translational start signal. In the examples, the

CC 3' end of pXT218 was attached to the signal DNA sequence of the DNA
CC fragment (CB6) for rat preproinsulin (see AAN10034)
XX
SQ Sequence 251 BP; 74 A; 45 C; 50 G; 82 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 1; Length 251;
Best Local Similarity 100.0%; Pred. No. 2.3e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAACAAATAG 60
Db AGGGTTATTGTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAACAAATAG 116
QY 61 GGGTTCCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 101
Db GGGTTCCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 75
RESULT 4
AAV31229/c
ID AAV31229 standard; DNA; 400 BP.
XX
AC AAV31229;
XX
DT 01-OCT-1998 (first entry)
XX
DE E. coli J96 pathogenicity island contig #43.
XX
KW PAI; pathogenicity island; uropathogenic E. coli detection; PAI IV; pheR;
KW PAI V; pheV; vaccine; protective immune response; ds.
XX
OS Escherichia coli.
XX
PN WO9822575-A2.
XX
PD 28-MAY-1998.
XX
PF 21-NOV-1997; 97WO-US021347.
PR 22-NOV-1996; 96US-0031626P.
PR 14-OCT-1997; 97US-0061953P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX (UYWI-) UNIV WISCONSIN.
XX
PI Dillon PJ, Choi GH, Welch RA;
XX WPI; 1998-312461/27.
XX
PT New isolated uropathogenic E. coli nucleotide sequences - used to develop
PT products for the detection of pathogenic E. coli and to elicit an immune
PT response to pathogenic E. coli.
PS Claim 21; Page 140-141; 250pp; English.
XX
CC This sequence represents a E. coli strain J96 contig containing
CC pathogenicity island (PAI) sequences, and represents a nucleic acid
CC molecule of the invention. PAIs are large fragments of DNA which comprise
CC pathogenicity determinants. The sequences of the invention are taken from
CC PAI IV and PAI V. PAI IV is located at approximately 64 min (near pheV)
CC on the E. coli chromosome and is greater than 170 kb. PAI V is located at
CC approximately 94 min (at pheR) on the E. coli chromosome and is
CC approximately 160 kb in size. Antibodies specific to the proteins encoded
CC by the PAI open reading frames of the invention can be used in kits to
CC detect uropathogenic E. coli. The proteins are used in vaccines to elicit
CC a protective immune response in an animal to the uropathogenic E. coli
CC strain J96
XX
SQ Sequence 400 BP; 106 A; 77 C; 91 G; 126 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 2; Length 400;
Best Local Similarity 100.0%; Pred. No. 2.5e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAACAAATAG 60
Db AGGGTTATTGTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAACAAATAG 106
QY 61 GGGTTCCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 101
Db GGGTTCCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 65
RESULT 5
AAN60624/c
ID AAN60624 standard; DNA; 456 BP.
XX
AC AAN60624;
XX
DT 25-MAR-2003 (revised)
DT 29-OCT-1991 (first entry)
XX
DE Plasmid pUG201 sequence encoding beta-urogastrone.
XX
KW Beta-lactamase signal peptide; pGH54; pGH55; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT 125..170
FT /*tag= a
FT RBS 200..203
FT /*tag= b
FT CDS 209..439
FT /*tag= c
FT sig_peptide 209..277
FT /*tag= d
FT /label= Beta-lactamase signal peptide
FT 278..436
FT /*tag= e
FT /label= Beta-urogastrone
XX
XX WO8603779-A.
XX
XX 03-JUL-1986.
XX
XX 19-DEC-1985; 85WO-JP000696.
XX
XX 21-DEC-1984; 84JP-00271206.
XX
XX (EART) EARTH CHEM CO LTD.
XX (OHGA/) OHGAI H.
XX
XX Ohgai H, Momota H, Kuwakura T, Kajifusa N, Kitazawa T, Oshiden K;
XX WPI; 1986-182911/28.
XX P-PSDB; AAP60678.
XX
XX Recombinant vector for polypeptide secretion - contains signal peptide
XX sequence directly bonded to peptide-coding sequence.
XX
XX Disclosure; Table 4; 79pp; Japanese.
XX
XX The plasmid produces secreted beta-urogastrone in a transformed
XX expression system. Similar plasmids may be constructed where the
XX secretion signal may be coupled with eg. somatostatin, insulin, growth
XX hormone, interferon, IL-2, gastric inhibitory peptide, influenza B SA,
XX epidermal growth factor and thymosine-beta4. (Updated on 25-MAR-2003 to
XX correct PA field.)
XX
SQ Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 1; Length 456;
Best Local Similarity 100.0%; Pred. No. 2.6e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGATTTAGAAAATAAACAATAG 60
Db 173 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGATTTAGAAAATAAACAATAG 114

Qy 61 GGGTTCCGCGCACATTTCCCGGAAAAGTGCCACCTGACGTC 101
Db 113 GGGTTCCGCGCACATTTCCCGGAAAAGTGCCACCTGACGTC 73

RESULT 6
AAN71080/c
ID AAN71080 standard; DNA; 456 BP.
XX
AC AAN71080;
XX
XX 25-MAR-2003 (revised)
DT 10-MAR-2003 (revised)
DT 13-MAY-1991 (first entry)
XX
DE Sequence encoding beta-urogastrone.
XX
XX pUGT 150s; beta-UG; ds.
XX
XX Escherichia coli.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH promoter 125..170
FT /*tag= a
FT CDS 209..439
FT /*tag= b
FT /*transl_except= (pos:434..436,aa:Arg)
XX
XX JP62190083-A.
XX
XX 20-AUG-1987.
XX
XX 14-FEB-1986; 86JP-00031415.
XX
XX 14-FEB-1986; 86JP-00031415.
XX
XX (EART ) EARTH SEIYAKU KK.
XX
XX WPI; 1987-273761/39.
XX
XX Expression vector contg. multiple information units is used to transform
PT host all for increased prodn. of polypeptides.
XX
XX Disclosure; Page 553; 34pp; Japanese.
XX
XX Sequence encodes beta-urogastrone under the control of a tac promoter.
CC The peptide may be expressed from plasmid pUGT 150s in a transformed
CC E.coli host. The plasmid may carry several separately expressing
CC sequences comprising a tac promoter, SD site, signal peptide, and coding
CC sequence, to produce beta-UG in high yield. (Updated on 10-MAR-2003 to
CC add missing OS field.) (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 1; Length 456;
Best Local Similarity 100.0%; Pred. No. 2.6e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGATTTAGAAAATAAACAATAG 60
Db 173 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGATTTAGAAAATAAACAATAG 114

Qy 61 GGGTTCCGCGCACATTTCCCGGAAAAGTGCCACCTGACGTC 101
Db 113 GGGTTCCGCGCACATTTCCCGGAAAAGTGCCACCTGACGTC 73

RESULT 7
AAN71080/c
ID AAN71080 standard; DNA; 456 BP.
XX
AC AAN71080;
XX
XX 25-MAR-2003 (revised)
DT 10-MAR-2003 (revised)
DT 13-MAY-1991 (first entry)
XX
DE Sequence encoding beta-urogastrone.
XX
XX pUGT 150s; beta-UG; ds.
XX
XX Escherichia coli.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH promoter 125..170
FT /*tag= a
FT CDS 209..439
FT /*tag= b
FT /*transl_except= (pos:434..436,aa:Arg)
XX
XX JP62190083-A.
XX
XX 20-AUG-1987.
XX
XX 14-FEB-1986; 86JP-00031415.
XX
XX 14-FEB-1986; 86JP-00031415.
XX
XX (EART ) EARTH SEIYAKU KK.
XX
XX WPI; 1987-273761/39.
XX
XX Expression vector contg. multiple information units is used to transform
PT host all for increased prodn. of polypeptides.
XX
XX Disclosure; Page 553; 34pp; Japanese.
XX
XX Sequence encodes beta-urogastrone under the control of a tac promoter.
CC The peptide may be expressed from plasmid pUGT 150s in a transformed
CC E.coli host. The plasmid may carry several separately expressing
CC sequences comprising a tac promoter, SD site, signal peptide, and coding
CC sequence, to produce beta-UG in high yield. (Updated on 10-MAR-2003 to
CC add missing OS field.) (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 1; Length 456;
Best Local Similarity 100.0%; Pred. No. 2.6e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGATTTAGAAAATAAACAATAG 60
Db 173 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGATTTAGAAAATAAACAATAG 114

Qy 61 GGGTTCCGCGCACATTTCCCGGAAAAGTGCCACCTGACGTC 101
Db 113 GGGTTCCGCGCACATTTCCCGGAAAAGTGCCACCTGACGTC 73

RESULT 8
AAN81765/c
ID AAN81765 standard; DNA; 456 BP.
XX
AC AAN81765;
XX
XX 25-MAR-2003 (revised)
DT 13-DEC-1990. (first entry)

```

```

AAN70833/c
ID AAN70833 standard; DNA; 456 BP.
XX
AC AAN70833;
XX
XX 25-MAR-2003 (revised)
DT 10-MAR-2003 (revised)
DT 18-JAN-1991 (first entry)
XX
DE Beta-urogastrone sequence.
XX
XX Tumour; inosine; DNA probe; ds.
XX
XX Unidentified.
XX
XX Key Location/Qualifiers
FH promoter 125..170
FT /*tag= b
FT RBS 200..204
FT /*tag= c
FT CDS 209..439
FT /*tag= a
FT sig_peptide 209..277
FT /*tag= d
XX
XX JP62244398-A.
XX
XX 24-OCT-1987.
XX
XX 16-APR-1986; 86JP-00087368.
XX
XX 16-APR-1986; 86JP-00087368.
XX
XX (SEKI ) SEKISUI CHEM IND CO LTD.
XX
XX WPI; 1987-339045/48.
XX
XX P-PSDB; AAP70505.
XX
XX Detection of DNA and/or RNA - by converting to single strand form and
PT using probe contg. labelled inosine deriv.
XX
XX Disclosure; Page 11; 11pp; Japanese.
XX
XX An example of a sequence detected by a probe consisting of polyinosine,
CC polydeoxyinosine, oligoinosine and/or oligodeoxyinosine is labeled. The
CC ssDNA and probe are hybridized and the existence of DNA in the product is
CC detected. It can be used to detect the presence of malignant tumour.
CC (Updated on 10-MAR-2003 to add missing OS field.) (Updated on 25-MAR-2003
CC to correct PA field.)
XX
SQ Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 1; Length 456;
Best Local Similarity 100.0%; Pred. No. 2.6e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGATTTAGAAAATAAACAATAG 60
Db 173 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGATTTAGAAAATAAACAATAG 114

Qy 61 GGGTTCCGCGCACATTTCCCGGAAAAGTGCCACCTGACGTC 101
Db 113 GGGTTCCGCGCACATTTCCCGGAAAAGTGCCACCTGACGTC 73

RESULT 8
AAN81765/c
ID AAN81765 standard; DNA; 456 BP.
XX
AC AAN81765;
XX
XX 25-MAR-2003 (revised)
DT 13-DEC-1990. (first entry)

```

```
XX Sequence of HB101 (pGH201) encoding new beta-urogastrone deriv. Met (21),
DE Arg (53).
XX
XX Gastric acid secretion; cell proliferation; hormone; ds.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX CDS 209..277
XX FT /*tag= a
XX FT 278..439
XX FT /*tag= b
XX FT /product= "New beta-urogastrone deriv."
XX
XX JP63012298-A.
XX
XX 19-JAN-1988.
XX
XX 30-JUN-1986; 86JP-00153783.
XX
XX 30-JUN-1986; 86JP-00153783.
XX
XX (EART ) EARTH SEIYAKU KK.
XX
XX WPI; 1988-054638/08.
XX
XX P-PSDB; AAP81349.
XX
XX New beta-urogastrone deriv. - has gastric acid secretion inhibition and
XX proliferation promotion activity.
XX
XX Disclosure; Page 685; 76pp; Japanese.
XX
XX The deriv. has various biological activities such as gastric acid
XX secretion inhibiting action, or cell proliferation promoting action. The
XX deriv. has the same biological or pharmacological activities as beta-
XX urogastrone. It is not susceptible to denaturation by oxidn. and is
XX chemically stable. Deriv. has resistance to proteolytic enzymes such as
XX pepsinase. (Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 101; DB 1; Length 456;
XX Best Local Similarity 100.0%; Pred. No. 2.6e-21;
XX Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAACAAATAG 60
XX 173 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAACAAATAG 114
XX
XX 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
XX 113 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 73
XX
XX RESULT 9
XX ABA90413/c
XX ID ABA90413 standard; DNA; 466 BP.
XX
XX AC ABA90413;
XX
XX 12-FEB-2002 (first entry)
XX
XX Drosophila cell cycle progression protein coding sequence #48.
XX
XX Antiproliferative; cytostatic; cardiant; immunosuppressive; meiosis;
XX antiinflammatory; antiposrotatic; dermatological; antifungal; mitosis;
XX antiparasitic; animalarial; antirheumatic; antiarthritic; cell division;
XX cell cycle progression protein; tumour; proliferative disorder;
XX cardiovascular; autoimmune; dermatological disorder; ds.
XX
XX Drosophila sp.
XX
XX PN WO200172774-A2.
XX
XX PD 04-OCT-2001.
XX
XX PF 23-MAR-2001; 2001WO-GB001297.
XX
XX PR 24-MAR-2000; 2000GB-00007268.
XX
XX PA (CYCL-) CYCLACEL LTD.
XX
XX PI Deak P, Glover DM, Midgley C;
XX
XX DR WPI; 2002-055132/07.
XX
XX Polynucleotides encoding cell cycle progression proteins, useful for
XX treating a tumor or a proliferative disorder.
XX
XX Claim 1; Page 99; 213pp; English.
XX
XX The present invention relates to Drosophila cell cycle progression
XX proteins (AAM47572-AAM47608) and their coding sequences (ABA90366-
XX ABA90520). The coding sequences and proteins are useful for identifying a
XX substance capable of affecting the function of the corresponding gene, a
XX inhibiting mitosis and/or meiosis. They can also be used in a method for
XX treating a tumour or proliferative disorder, cardiovascular disorders
XX (such as restenosis and cardiomyopathy), autoimmune disorders such as
XX glomerulonephritis and rheumatoid arthritis), dermatological disorders
XX (such as psoriasis), antiinflammatory, antifungal and antiparasitic
XX disorders (such as malaria)
XX
XX Sequence 466 BP; 135 A; 87 C; 93 G; 150 T; 0 U; 1 Other;
XX
XX Query Match 100.0%; Score 101; DB 6; Length 466;
XX Best Local Similarity 100.0%; Pred. No. 2.6e-21;
XX Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAACAAATAG 60
XX 280 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAACAAATAG 221
XX
XX 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
XX 220 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 180
XX
XX RESULT 10
XX AAX21173/c
XX ID AAX21173 standard; DNA; 487 BP.
XX
XX AC AAX21173;
XX
XX 05-MAY-1999 (first entry)
XX
XX Polynucleotide sequence from the genome of Treponema pallidum.
XX
XX Treponema pallidum infection; syphilis; Borrelia infection; animal;
XX enzyme production; ds.
XX
XX Treponema pallidum.
XX
XX WO9859034-A2.
XX
XX 30-DEC-1998.
XX
XX 23-JUN-1998; 98WO-US013041.
XX
XX 24-JUN-1997; 97US-0050667P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Fraser CM;
XX
```

DR WPI; 1999-081273/07.

XX New isolated *Treponema pallidum* nucleic acids - used to develop products

PT for the detection, diagnosis, characterisation, prevention and therapy of

PT *T. pallidum* infections, particularly syphilis.

XX

PS Claim 1; Page 1106; 1150pp; English.

XX

CC AAX20500-21243 represent polynucleotide sequences from the genome of

CC *Treponema pallidum*. The sequences can be used for detection, diagnosis,

CC characterisation, prevention and therapy for *T. pallidum* infections,

CC particularly syphilis. They can also be used for detecting diseases

CC related to *Borrelia* infections in animals, and for the production of

CC biosynthetic products such as enzymes

XX

XX Sequence 487 BP; 125 A; 127 C; 113 G; 121 T; 0 U; 1 Other;

SQ

Query Match 100.0%; Score 101; DB 2; Length 487;

Best Local Similarity 100.0%; Pred. No. 2.6e-21;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTGAGAAAATAACAAATAG 60

DB 323 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTGAGAAAATAACAAATAG 264

OY 61 GGGTTCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 101

DB 263 GGGTTCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 223

RESULT 11

AAX21149/c

ID AAX21149 standard; DNA; 535 BP.

XX

AC AAX21149;

XX

XX 05-MAY-1999 (first entry)

XX

XX Polynucleotide sequence from the genome of *Treponema pallidum*.

DE

XX *Treponema pallidum* infection; syphilis; *Borrelia* infection; animal;

KW enzyme production; ds.

XX

OS *Treponema pallidum*.

XX

XX WO9859034-A2.

PN

XX 30-DEC-1998.

PD

XX 23-JUN-1998; 98WO-US013041.

PF

XX 24-JUN-1997; 97US-0050667P.

PR

XX (HUMA-) HUMAN GENOME SCI INC.

PA

XX Fraser CM;

PI

XX WPI; 1999-081273/07.

DR

XX New isolated *Treponema pallidum* nucleic acids - used to develop products

PT for the detection, diagnosis, characterisation, prevention and therapy of

PT *T. pallidum* infections, particularly syphilis.

XX

PS Claim 1; Page 1093; 1150pp; English.

XX

CC AAX20500-21243 represent polynucleotide sequences from the genome of

CC *Treponema pallidum*. The sequences can be used for detection, diagnosis,

CC characterisation, prevention and therapy for *T. pallidum* infections,

CC particularly syphilis. They can also be used for detecting diseases

CC related to *Borrelia* infections in animals, and for the production of

CC biosynthetic products such as enzymes

XX

XX Sequence 535 BP; 145 A; 108 C; 122 G; 155 T; 0 U; 5 Other;

SQ

Query Match 100.0%; Score 101; DB 6; Length 573;

Best Local Similarity 100.0%; Pred. No. 2.7e-21;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTGAGAAAATAACAAATAG 60

DB 355 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTGAGAAAATAACAAATAG 296

Query Match 100.0%; Score 101; DB 2; Length 535;

Best Local Similarity 100.0%; Pred. No. 2.7e-21;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTGAGAAAATAACAAATAG 60

DB 158 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTGAGAAAATAACAAATAG 99

OY 61 GGGTTCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 101

DB 98 GGGTTCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 58

RESULT 12

ABA90456/c

ID ABA90456 standard; DNA; 573 BP.

XX

AC ABA90456;

XX

XX 12-FEB-2002 (first entry)

DT

XX *Drosophila* cell cycle progression protein coding sequence #31.

DE

XX Antiproliferative; cytostatic; cardiant; immunosuppressive; meiosis;

KW antiinflammatory; antipsoriatic; dermatological; antifungal; mitosis;

KW antiparasitic; antimalarial; antirheumatic; antiarthritic; cell division;

KW cell cycle progression protein; tumour; proliferative disorder;

KW cardiovascular; autoimmune; dermatological disorder; ds.

XX

OS *Drosophila* sp.

XX

XX WO200172774-A2.

PN

XX 04-OCT-2001.

PD

XX 23-MAR-2001; 2001WO-GB001297.

PF

XX 24-MAR-2000; 2000GB-00007268.

PR

XX (CYCL-) CYCLACEL LTD.

PA

XX Deak P, Glover DM, Midgley C;

PI

XX WPI; 2002-055132/07.

DR

XX Polynucleotides encoding cell cycle progression proteins, useful for

PT treating a tumor or a proliferative disorder.

PT

XX Claim 1; Page 144; 213pp; English.

PS

XX The present invention relates to *Drosophila* cell cycle progression

CC proteins (AAM47572-AAM47608) and their coding sequences (ABA90366-

CC ABA90520). The coding sequences and proteins are useful for identifying a

CC substance capable of affecting the function of the corresponding gene, a

CC substance capable of inhibiting the cell division cycle, or a method of

CC inhibiting mitosis and/or meiosis. They can also be used in a method for

CC treating a tumour or proliferative disorder, cardiovascular disorders

CC (such as restenosis and cardiomyopathy), autoimmune disorders such as

CC (glomerulonephritis and rheumatoid arthritis), dermatological disorders

CC (such as psoriasis), antiinflammatory, antifungal and antiparasitic

CC disorders (such as malaria)

XX

XX Sequence 573 BP; 154 A; 118 C; 116 G; 184 T; 0 U; 1 Other;

SQ

Query Match 100.0%; Score 101; DB 6; Length 573;

Best Local Similarity 100.0%; Pred. No. 2.7e-21;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTGAGAAAATAACAAATAG 60

DB 355 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTGAGAAAATAACAAATAG 296

QY 61 GGGTTCGCGCACATTTCCCGAAAAAGTGCACCTGACGTC 101
|||||
Db 295 GGGTTCGCGCACATTTCCCGAAAAAGTGCACCTGACGTC 255

RESULT 13
ADH58311
ID ADH58311 standard; DNA; 605 BP.

AC ADH58311;
XX
XX 25-MAR-2004 (first entry)
DT
DE Electropherogram of a DNA sequencing reaction using E154 & T422.
XX ds; primer library; extendable oligos; EO; ligation chain reaction; LCR;
XX rolling circle amplification; strand displacement amplification;
KW isothermal DNA amplification; biotechnology; agriculture;
KW medical research; pUC19 plasmid.

XX Synthetic.
OS Escherichia coli.
XX WO2003093500-A1.

XX 13-NOV-2003.
XX 24-DEC-2002; 2002WO-AU001763.
XX 01-MAY-2002; 2002AU-00002045.

XX (NUCL-) NUCLEICS PTY LTD.

XX Tillett D, Thomas T;

XX WPI; 2004-053046/05.

XX Increasing the affinity of an extendable oligonucleotide (EO) for a target nucleic acid, for providing primers having improved specificity, comprises hybridization of the EO to a template oligonucleotide (TO) and extension of the EO.

PS Example 10; Fig 23; 85pp; English.

XX This invention relates to a novel method for the optimisation of primer libraries. Specifically, it refers to increasing the affinity of short oligonucleotide primers, also known as extendable oligos (EOs), for their template sequences. The present invention describes improved methods for sequencing and the linear and exponential amplification of DNA that can be useful for PCR, RT-PCR, ligation chain reaction (LCR), rolling circle amplification, strand displacement amplification and isothermal DNA amplification. Accordingly, these extendable oligos with improved specificity and affinity are particularly important in fields ranging from biotechnology and agriculture to medical research. This polynucleotide sequence is the electropherogram of a DNA sequencing reaction that used the pUC19 plasmid and E154/T422 oligos, used in an exemplification of the invention.

XX Sequence 605 BP; 159 A; 133 C; 147 G; 148 T; 0 U; 18 Other;

Query Match 100.0%; Score 101; DB 12; Length 605;
Best Local Similarity 100.0%; Pred. No. 2.8e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATTGTCATGACGGATACATATTTGAATGTATTAGAAAAATAACAATAAG 60
|||||
Db 259 AGGGTTATTGTCATGACGGATACATATTTGAATGTATTAGAAAAATAACAATAAG 318

QY 61 GGGTTCGCGCACATTTCCCGAAAAAGTGCACCTGACGTC 101
|||||
Db 319 GGGTTCGCGCACATTTCCCGAAAAAGTGCACCTGACGTC 359

RESULT 14
AAS30560/c
ID AAS30560 standard; DNA; 776 BP.
XX
AC AAS30560;
XX
XX 21-NOV-2001 (first entry)
DT
DE DNA encoding novel prostate gland antigen, Seq ID No 418.
XX
XX Human; neutrotropic; neuroprotective; cytostatic; antiparkinsonian;
KW antineoplastic; dermatological; immunosuppressive; antiinflammatory;
KW antiarthritic; antirheumatic; virucide; hepatotropic; nephrotropic;
KW osteopathic; prostate gland; prostatitis; adenocarcinoma; hair loss;
KW prostatic; malacoplakia; adenocarcinoma; benign prostatic hypertrophy;
KW hyperplasia; carcinoma; prostate neoplastic disorder; skin aging;
KW reproductive system disorder; autoimmune disorder; urinary system;
KW systemic lupus erythematosus; rheumatoid arthritis; cardiovascular;
KW blood-related disorder; hyperproliferative disorder; respiratory;
KW neurological disorder; endocrine disorder; inflammatory disorder;
KW liver disorder; wound healing; food preservative; ds.
XX
XX Homo sapiens.
OS
XX WO200155447-A1.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001330.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.

XX AAS27819;
XX 07-NOV-2001 (first entry)
XX DNA encoding novel signal transduction pathway protein, Seq ID 1479.
XX Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
KW antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;
KW immune system disorder; rheumatoid arthritis; inflammatory condition;
KW organ transplant rejection; infection; hepatitis C; blood disorder;
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
KW reproductive system; gastrointestinal; liver disorder; AIDS; ds;
XX acquired immune deficiency syndrome.
OS Homo sapiens.
XX WO200154733-A1.
XX 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US001312.
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0198974P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0203467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241836P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 05:15:57 ; Search time 961.667 Seconds
(without alignments)
3997.736 Million cell updates/sec

Title: US-09-482-682-43_COPY_8283_8383
Perfect score: 101
Sequence: 1 agggttattgtctcatgagc.....gaaaagtccacctgacgtc 101

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	101	100.0	300	4	BM078095
C 2	101	100.0	300	5	BU963956
C 3	101	100.0	300	5	BU964094
C 4	101	100.0	309	9	FR0099140
C 5	101	100.0	391	1	AL597149
C 6	101	100.0	414	9	CC819240
C 7	101	100.0	417	4	BJ684174
C 8	101	100.0	491	9	CC819923
C 9	101	100.0	495	4	BI805285
C 10	101	100.0	495	9	CC818374
C 11	101	100.0	496	9	CC818523
C 12	101	100.0	503	9	CC819854
C 13	101	100.0	515	9	CC817752
C 14	101	100.0	518	9	CC817128
C 15	101	100.0	519	9	CC817162
C 16	101	100.0	519	9	CC817796
C 17	101	100.0	521	9	CC819067
C 18	101	100.0	533	9	CC819841
C 19	101	100.0	542	9	CC816892
C 20	101	100.0	550	7	CF766622
C 21	101	100.0	551	9	CC818905
C 22	101	100.0	554	9	CC819058
C 23	101	100.0	563	9	CC819270
C 24	101	100.0	566	9	CC816848

C 25	101	100.0	566	9	CC820024
C 26	101	100.0	567	9	CC817070
C 27	101	100.0	568	9	CC818640
C 28	101	100.0	571	9	CC816954
C 29	101	100.0	571	9	CC818423
C 30	101	100.0	580	9	CC819098
C 31	101	100.0	582	1	AL694813
C 32	101	100.0	583	9	CC817633
C 33	101	100.0	583	9	CC818436
C 34	101	100.0	586	9	CC816883
C 35	101	100.0	588	9	CC817788
C 36	101	100.0	588	9	CC818340
C 37	101	100.0	589	9	CC817595
C 38	101	100.0	590	9	CC819754
C 39	101	100.0	592	9	CC817679
C 40	101	100.0	592	9	CC818508
C 41	101	100.0	593	9	CC816942
C 42	101	100.0	593	9	CC817699
C 43	101	100.0	594	9	CC818287
C 44	101	100.0	594	9	CC818422
C 45	101	100.0	595	9	CC816929

ALIGNMENTS

RESULT 1
BM078095/c
LOCUS BM078095 300 bp mRNA linear EST 30-NOV-2001
DEFINITION 83374 Hebeloma cylindrosporum functional cDNA library Hebeloma cylindrosporum cDNA 5', mRNA sequence.
ACCESSION BM078095
VERSION BM078095.1 GI:17157967
KEYWORDS EST.
SOURCE Hebeloma cylindrosporum
ORGANISM Hebeloma cylindrosporum
REFERENCE 1 (bases 1 to 300)
AUTHORS Wipf,D., Benjdia,M., Tegeder,M. and Frommer,W.B.
TITLE Construction of a functional cDNA library from the ectomycorrhizal fungus Hebeloma cylindrosporum
JOURNAL Unpublished (2001)
COMMENT Contact: Wipf D
ZMBP - Center for Molecular Biology of Plants
University of Tuebingen
Auf der Morgenstelle 1, 72076 Tuebingen, Germany
Tel: 49 7071 2976160
Fax: 49 7071 293287
Email: daniel.wipf@zmbp.uni-tuebingen.de
PCR Primers
FORWARD: pDR196 5' primer (PMA 5')
HIGH quality sequence stop: 300
POLYA-No.

FEATURES
source
1..300
/organism="Hebeloma cylindrosporum"
/mol_type="mRNA"
/strain="H1"
/db_xref="taxon:76867"
/tissue_type="Mycelia"
/lab_host="E. coli XLI-Blue"
/clone_lib="Hebeloma cylindrosporum functional cDNA library"
/note="vector: pDR 196 (unpublished); Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Query Match 100.0%; Score 101; DB 4; Length 300;
Best Local Similarity 100.0%; Pred. No. 8.1e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGGTTATTGCTCATGCGGATACATATTGTAATGTTAGAAAAATAACAAATAG 60

```

|||||
174 AGGGTTATTGCTCATGACGGATACATATTGAATTTAGAAAAATAACAAATAG 115
|||||

Qy 61 GGGTTCGCGCACATTTCCCGGAAAGTGCCACCTGAGTC 101
|||||
Db 114 GGGTTCGCGCACATTTCCCGGAAAGTGCCACCTGAGTC 74
|||||

RESULT 2
BU963956/c 300 bp mRNA linear EST 13-NOV-2002
LOCUS
DEFINITION ESP88 Hebeloma cylindrosporum functional cDNA library Hebeloma
cylindrosporum cDNA, mRNA sequence.
ACCESSION BU963956
VERSION BU963956.1 GI:24204753
KEYWORDS EST.
SOURCE Hebeloma cylindrosporum
ORGANISM Hebeloma cylindrosporum
Eukaryota; Fungi; Basidiomycetes; Hymenomycetes; Homobasidiomycetes;
Agaricales; Cortinariaceae; Hebeloma.
REFERENCE 1 (bases 1 to 300)
AUTHORS Wipf,D., Benjdia,M., Rikirsch,E., Zimmermann,S., Tegeder,M. and
Frommer,W.B.
TITLE A Hebeloma cylindrosporum expression cDNA library for suppression
cloning in yeast mutants, complementation of a yeast his4 mutant
and EST analysis
JOURNAL Unpublished (2002)
COMMENT ZMBP - Center for Molecular Biology of Plants
University of Tuebingen
Auf der Morgenstelle 1, 72076 Tuebingen, Germany
Tel: 49 7071 2976160
Fax: 49 7071 293287
Email: daniel.wipf@zmbp.uni-tuebingen.de
homolog to Enterobacteria phage f1 ; ampicillinase (1e-10) .

FEATURES
source
1..300
/organism="Hebeloma cylindrosporum"
/mol_type="mRNA"
/strain="H1"
/db_xref="taxon:76867"
/tissue_type="Mycelia"
/lab_hosts="E. coli XL1-Blue"
/clone_lib="Hebeloma cylindrosporum functional cDNA
library"
/notes="Vector: pDR 196 (unpublished); Site_1: EcoRI;
Site_2: XhoI"

ORIGIN
Query Match 100.0%; Score 101; DB 5; Length 300;
Best Local Similarity 100.0%; Pred. No. 8.1e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGACGGATACATATTGAATTTAGAAAAATAACAAATAG 60
|||||
Db 170 AGGGTTATTGCTCATGACGGATACATATTGAATTTAGAAAAATAACAAATAG 111
|||||

Qy 61 GGGTTCGCGCACATTTCCCGGAAAGTGCCACCTGAGTC 101
|||||
Db 110 GGGTTCGCGCACATTTCCCGGAAAGTGCCACCTGAGTC 70
|||||

RESULT 4
FR0009140 309 bp DNA linear GSS 25-FEB-2004
LOCUS
DEFINITION F.rubripes GSS sequence, clone 010H20ac4, genomic survey sequence.
ACCESSION AL000426
VERSION AL000426.1 GI:2438278
KEYWORDS GSS; genome survey sequence.
SOURCE Takifugu rubripes (Fugu rubripes)
ORGANISM Takifugu rubripes
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontoidea; Tetraodontidae; Takifugu.

REFERENCE 1
AUTHORS Elgar,G., Clark,M.S., Meek,S., Smith,S., Warner,S., Edwards,Y.J.,
Bouchireb,N., Cottage,A., Yeo,G.S., Umrانيا,Y., Williams,G. and
Brenner,S.
TITLE Generation and analysis of 25 Mb of genomic DNA from the pufferfish
Fugu rubripes by sequence scanning
JOURNAL Genome Res. 9 (10), 960-971 (1999)
MEDLINE 99455097
PUBMED 10523524
REFERENCE 2 (bases 1 to 309)
AUTHORS Elgar,G., Clark,M., Smith,S., Meek,S., Warner,S., Umrانيا,Y.,
Williams,G. and Brenner,S.
TITLE Direct Submission
JOURNAL Submitted (09-SEP-1997) MRC Human Genome Mapping Project Resource
Centre Hinxton, Cambridge, CB10 1SB. Email: biohelp@hgmrc.ac.uk
COMMENT Vector: pBluescript II KS
V_type: phagemid

```

```

Agaricales; Cortinariaceae; Hebeloma.
REFERENCE 1 (bases 1 to 300)
AUTHORS Wipf,D., Benjdia,M., Rikirsch,E., Zimmermann,S., Tegeder,M. and
Frommer,W.B.
TITLE A Hebeloma cylindrosporum expression cDNA library for suppression
cloning in yeast mutants, complementation of a yeast his4 mutant
and EST analysis
JOURNAL Unpublished (2002)
COMMENT ZMBP - Center for Molecular Biology of Plants
University of Tuebingen
Auf der Morgenstelle 1, 72076 Tuebingen, Germany
Tel: 49 7071 2976160
Fax: 49 7071 293287
Email: daniel.wipf@zmbp.uni-tuebingen.de
no homology below 1e-10.

FEATURES
source
1..300
/organism="Hebeloma cylindrosporum"
/mol_type="mRNA"
/strain="H1"
/db_xref="taxon:76867"
/tissue_type="Mycelia"
/lab_hosts="E. coli XL1-Blue"
/clone_lib="Hebeloma cylindrosporum functional cDNA
library"
/notes="Vector: pDR 196 (unpublished); Site_1: EcoRI;
Site_2: XhoI"

ORIGIN
Query Match 100.0%; Score 101; DB 5; Length 300;
Best Local Similarity 100.0%; Pred. No. 8.1e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGACGGATACATATTGAATTTAGAAAAATAACAAATAG 60
|||||
Db 170 AGGGTTATTGCTCATGACGGATACATATTGAATTTAGAAAAATAACAAATAG 111
|||||

Qy 61 GGGTTCGCGCACATTTCCCGGAAAGTGCCACCTGAGTC 101
|||||
Db 110 GGGTTCGCGCACATTTCCCGGAAAGTGCCACCTGAGTC 70
|||||

RESULT 4
FR0009140 309 bp DNA linear GSS 25-FEB-2004
LOCUS
DEFINITION F.rubripes GSS sequence, clone 010H20ac4, genomic survey sequence.
ACCESSION AL000426
VERSION AL000426.1 GI:2438278
KEYWORDS GSS; genome survey sequence.
SOURCE Takifugu rubripes (Fugu rubripes)
ORGANISM Takifugu rubripes
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontoidea; Tetraodontidae; Takifugu.

REFERENCE 1
AUTHORS Elgar,G., Clark,M.S., Meek,S., Smith,S., Warner,S., Edwards,Y.J.,
Bouchireb,N., Cottage,A., Yeo,G.S., Umrانيا,Y., Williams,G. and
Brenner,S.
TITLE Generation and analysis of 25 Mb of genomic DNA from the pufferfish
Fugu rubripes by sequence scanning
JOURNAL Genome Res. 9 (10), 960-971 (1999)
MEDLINE 99455097
PUBMED 10523524
REFERENCE 2 (bases 1 to 309)
AUTHORS Elgar,G., Clark,M., Smith,S., Meek,S., Warner,S., Umrانيا,Y.,
Williams,G. and Brenner,S.
TITLE Direct Submission
JOURNAL Submitted (09-SEP-1997) MRC Human Genome Mapping Project Resource
Centre Hinxton, Cambridge, CB10 1SB. Email: biohelp@hgmrc.ac.uk
COMMENT Vector: pBluescript II KS
V_type: phagemid

```

```
PRIMER: KS
DESCR:
One pass dye-terminator sequencing of cosmid cloned genomic
sequence.
FEATURES
    source
        Location/Qualifiers
            1..309
                /organism="Takifugu rubripes"
                /mol_type="genomic DNA"
                /db_xref="taxon:31033"
                /clone="01OH20aC4"
                /clone_lib="cosmid 01OH20"
ORIGIN
    Query Match      100.0%; Score 101; DB 9; Length 309;
    Best Local Similarity 100.0%; Pred. No. 8.2e-19;
    Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAACAAATAG 60
Db 39 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAACAAATAG 98
Qy 61 GGGTTCGGCGACATTTCCCGGAAAAGTGCACCTGACGTC 101
Db 99 GGGTTCGGCGACATTTCCCGGAAAAGTGCACCTGACGTC 139
RESULT 5
AL597149
LOCUS
DEFINITION
DKFZp313J1611_r1 313 (synonym: hlcc2) Homo sapiens cDNA clone
AL597149
ACCESSION
VERSION
KEYWORDS
SOURCE
    Homo sapiens (human)
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
    1 (bases 1 to 391)
    Koehrer, K., Beyer, A., Mewes, W., Weil, B. and Wiemann, S.
    EST (Koehrer, K., Beyer, A., Mewes, H.W., Weil, B. and Wiemann, S.)
    Unpublished (1999)
    Contact: MIPS
    MIPS
    Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
    This is the 5' sequence of the clone insert
    Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
    Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
    sequenced by BMF2 (Biomedical Research Center at the Charite,
    Berlin/Germany) within the cDNA sequencing consortium of the German
    Genome Project.
    No sl sequence available.
    This clone (DKFZp313J1611) is available at the RZPD in Berlin.
    Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
    Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
    source
        Location/Qualifiers
            1..391
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="DKFZp313J1611"
                /dev_stage="adult"
                /lab_host="DH10B"
                /clone_lib="313 (synonym: hlcc2)"
                /note="Vector: pTriplex2; Site_1: SfIa; Site_2: SfIIB;
                cDNA-collection"
ORIGIN
    Query Match      100.0%; Score 101; DB 1; Length 391;
    Best Local Similarity 100.0%; Pred. No. 8.3e-19;
    Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAACAAATAG 60
Qy 61 GGGTTCGGCGACATTTCCCGGAAAAGTGCACCTGACGTC 101
Db 354 GGGTTCGGCGACATTTCCCGGAAAAGTGCACCTGACGTC 314
Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAACAAATAG 355
Db 414 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAACAAATAG 355
PRIMER: KS
DESCR:
One pass dye-terminator sequencing of cosmid cloned genomic
sequence.
FEATURES
    source
        Location/Qualifiers
            1..309
                /organism="Takifugu rubripes"
                /mol_type="genomic DNA"
                /db_xref="taxon:31033"
                /clone="01OH20aC4"
                /clone_lib="cosmid 01OH20"
ORIGIN
    Query Match      100.0%; Score 101; DB 9; Length 309;
    Best Local Similarity 100.0%; Pred. No. 8.2e-19;
    Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAACAAATAG 60
Db 39 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAACAAATAG 98
Qy 61 GGGTTCGGCGACATTTCCCGGAAAAGTGCACCTGACGTC 101
Db 99 GGGTTCGGCGACATTTCCCGGAAAAGTGCACCTGACGTC 139
RESULT 6
CC819240
LOCUS
DEFINITION
100005D19R Oxytricha plasmid UUGC10 library Sterkiella
histriomuscorum genomic clone UUGC100005D19 R, genomic survey
sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
    Sterkiella histriomuscorum (Oxytricha trifallax)
    Sterkiella histriomuscorum
    Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
    Stichotrichida; Oxytrichidae; Sterkiella.
REFERENCE
    1 (bases 1 to 414)
    Dunn, D., Doak, T., Herrick, G. and Weiss, R.
    Paired end reads from plasmid inserts of Oxytricha trifallax
    macronuclear chromosomes
    Unpublished (2003)
    Contact: Robert B. Weiss
    University of Utah Genome Center
    University of Utah
    Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
    84112, USA
    Tel: 801 585 5606
    Fax: 801 585 7177
    Email: ddunn@genetics.utah.edu
    Plate: 0005 row: D column: 19
    Seq primer: CACACAGGAACAGCTATGACC
    Class: plasmid ends
    High quality sequence stop: 414.
FEATURES
    source
        Location/Qualifiers
            1..414
                /organism="Sterkiella histriomuscorum"
                /mol_type="genomic DNA"
                /db_xref="taxon:94289"
                /clone="UUGC100005D19"
                /lab_host="E. Coli strain XL10-Gold, Ti-resistant, F-"
                /clone_lib="Oxytricha plasmid UUGC10 library"
                /note="Vector: PWD42nv; Purified macronuclear chromosomal
                DNA from Oxytricha trifallax was blunt end-repaired with
                T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
                oligonucleotides were ligated to the blunt ends in high
                molar excess. Vector DNA was prepared from a derivative of
                PWD42 (gil4732114|gb|AF129072.1), a copy-number inducible
                derivative of plasmid RI. The vector was ligated with
                adaptors complementary to the insert adaptors and
                purified. The sheared, adapted mouse DNA was annealed to
                adapted vector DNA, and transformed into
                chemically-competent E. Coli XL10-Gold (Stratagene) cells
                and selected for ampicillin resistance."
ORIGIN
    Query Match      100.0%; Score 101; DB 9; Length 414;
    Best Local Similarity 100.0%; Pred. No. 8.3e-19;
    Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAACAAATAG 60
Db 414 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAACAAATAG 355
Qy 61 GGGTTCGGCGACATTTCCCGGAAAAGTGCACCTGACGTC 101
Db 354 GGGTTCGGCGACATTTCCCGGAAAAGTGCACCTGACGTC 314
```

```

RESULT 7
BJ684174/c
LOCUS
DEFINITION BJ684174 HCEST library Haplochromis chilotes cDNA clone no90c12,
            mRNA sequence.
ACCESSION BJ684174
VERSION BJ684174.1 GI:46527295
KEYWORDS EST.
SOURCE
  ORGANISM Haplochromis chilotes
            Euplochromis chilotes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Acanthopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Actinopterygii; Acanthopterygii; Percormorpha; Perciformes;
            Labroidae; Cichlidae; Haplochromis.
REFERENCE 1 (bases 1 to 417)
AUTHORS Watanabe,M., Kobayashi,N., Shin-i,T., Kohara,Y. and Okada,N.
TITLE Orf sequences of cichlid in Lake Victoria are essentially same
JOURNAL Unpublished (2004)
COMMENT Contact: Tadaeu Shin-i
          Center For Genetic Resource Information
          National Institute of Genetics
          1111 Yata, Mishima, Shizuoka 411-8540, Japan
          Tel: 81-559-81-6856
          Fax: 81-559-81-6855
          Email: tshini@gene.nig.ac.jp.
FEATURES
  source
    Location/Qualifiers
      1..417
        /organism="Haplochromis chilotes"
        /mol_type="mRNA"
        /db_xref="taxon:257977"
        /clone="no90c12"
        /tissue_type="jaw"
        /dev_stage="varied"
        /clone_lib="HCEST library"

ORIGIN
Query Match 100.0%; Score 101; DB 4; Length 417;
Best Local Similarity 100.0%; Pred. No. 8.3e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGTTATTGCTCATGAGCGGATACATATTGAATGTTAGAAAAATAACAATAG 60
Db 129 AGGTTATTGCTCATGAGCGGATACATATTGAATGTTAGAAAAATAACAATAG 70

Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 101
Db 69 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 29

RESULT 8
CC819923/c
LOCUS
DEFINITION CC819923 Oxytricha plasmid UUGC10006J13 R, genomic survey
            histriomuscorum genomic clone UUGC10006J13 R, genomic survey
            sequence.
ACCESSION CC819923
VERSION CC819923.1 GI:32900671
KEYWORDS GSS.
SOURCE
  ORGANISM Sterkiella histriomuscorum (Oxytricha trifallax)
            Sterkiella histriomuscorum
            Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
            Stichotrichida; Oxytrichidae; Sterkiella.
REFERENCE 1 (bases 1 to 491)
AUTHORS Dunn,D., Doak,T., Herrick,G. and Weiss,R.
TITLE Paired end reads from plasmid inserts of Oxytricha trifallax
JOURNAL macronuclear chromosomes
COMMENT Unpublished (2003)
          Contact: Robert B. Weiss
          University of Utah Genome Center
          Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
          84112, USA
          Tel: 801 585 5606

FEATURES
  source
    Location/Qualifiers
      1..491
        /organism="Sterkiella histriomuscorum"
        /mol_type="genomic DNA"
        /db_xref="taxon:94289"
        /clone="UUGC10006J13"
        /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
        /clone_lib="Oxytricha plasmid UUGC10 library"
        /notes="Vector: pWD42nv; Purified macronuclear chromosomal
        DNA from Oxytricha trifallax was blunt end-repaired with
        T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
        oligonucleotides were ligated to the blunt ends in high
        molar excess. Vector DNA was prepared from a derivative of
        pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
        derivative of plasmid R1. The vector was ligated with
        adaptors complementary to the insert adaptors and
        purified. The sheared, adapted mouse DNA was annealed to
        adapted vector DNA, and transformed into
        chemically-competent E. Coli XL10-Gold (Stratagene) cells
        and selected for ampicillin resistance."

```

```

FEATURES
  source
    Location/Qualifiers
      1..491
        /organism="Sterkiella histriomuscorum"
        /mol_type="genomic DNA"
        /db_xref="taxon:94289"
        /clone="UUGC10006J13"
        /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
        /clone_lib="Oxytricha plasmid UUGC10 library"
        /notes="Vector: pWD42nv; Purified macronuclear chromosomal
        DNA from Oxytricha trifallax was blunt end-repaired with
        T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
        oligonucleotides were ligated to the blunt ends in high
        molar excess. Vector DNA was prepared from a derivative of
        pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
        derivative of plasmid R1. The vector was ligated with
        adaptors complementary to the insert adaptors and
        purified. The sheared, adapted mouse DNA was annealed to
        adapted vector DNA, and transformed into
        chemically-competent E. Coli XL10-Gold (Stratagene) cells
        and selected for ampicillin resistance."

ORIGIN
Query Match 100.0%; Score 101; DB 9; Length 491;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGTTATTGCTCATGAGCGGATACATATTGAATGTTAGAAAAATAACAATAG 60
Db 412 AGGTTATTGCTCATGAGCGGATACATATTGAATGTTAGAAAAATAACAATAG 353

Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 101
Db 352 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 312

RESULT 9
BI805285
LOCUS
DEFINITION BI805285 Stem library from Oryza sativa (3-5 leaf stage) Oryza
            sativa cDNA clone S035A01, mRNA sequence.
ACCESSION BI805285
VERSION BI805285.1 GI:15852489
KEYWORDS EST.
SOURCE
  ORGANISM Oryza sativa
            Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 495)
AUTHORS Dong,H.T., Li,D.B., Zhuang,X.F., Dai,C.G., Sun,L.X., Pei,Y.X.,
            Wu,H.F., Jiang,Y.X., Yu,F.C., Gao,Q.K. and Lou,Y.C.
TITLE A Gene Expression Screen in Oryza sativa
JOURNAL Unpublished (2001)
COMMENT Contact: Haitao Dong, Debao Li
            Bioinformatics and Gene Network Research Group
            Zhejiang University
            Kaixuan Road 268#, Hangzhou, Zhejiang, P.R.China
            Tel: 0086-571-8692051
            Fax: 0086-571-86961525
            Email: webmaster@estarray.org, URL: http://www.estarray.org
            Seq primer: M13 forward primer.
FEATURES
  source
    Location/Qualifiers
      1..495
        /organism="Oryza sativa"
        /mol_type="mRNA"
        /db_xref="taxon:4530"
        /clone="S035A01"

```



```
/tissue_type="Stem"
/dev_stage="3-5 leaf stage"
/clone_lib="Stem library from Oryza sativa (3-5 leaf
stage)"
/note="Vector: pSport2"

ORIGIN

Query Match
Best Local Similarity 100.0%; Score 101; DB 4; Length 495;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTAGAAAATAAACAATAG 60
|
Db 62 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTAGAAAATAAACAATAG 121
|
Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 101
|
Db 122 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 162
|

RESULT 10
CC818374/c
LOCUS
DEFINITION
CC818374 495 bp DNA linear GSS 17-JUL-2003
100004807R Oxytricha plasmid UUGC10 library Sterkiella
histriomuscorum genomic clone UUGC100004807 R, genomic survey
sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Sterkiella histriomuscorum (Oxytricha trifallax)
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
Stichotrichida; Oxytrichidae; Sterkiella.
REFERENCE
AUTHORS
1 (bases 1 to 495)
Dunn,D., Doak,T., Herrick,G. and Weiss,R.
Paired end reads from plasmid inserts of Oxytricha trifallax
macronuclear chromosomes
Unpublished (2003)
JOURNAL
COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Plate: 0004 row: B column: 07
Seq primer: CACACAGGAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 495.
Location/Qualifiers
1. 495
/organism="Sterkiella histriomuscorum"
/mol_type="genomic DNA"
/db_xref="taxon:94289"
/clone="UUGC100004807"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Oxytricha plasmid UUGC10 library"
/note="Vector: PWD42nv; Purified macronuclear chromosomal
DNA from Oxytricha trifallax was blunt end-repaired with
T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
oligonucleotides were ligated to the blunt ends in high
molar excess. Vector DNA was prepared from a derivative of
PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
derivative of plasmid R1. The vector was ligated with
adaptors complementary to the insert adaptors and
purified. The sheared, adaptor mouse DNA was annealed to
adaptor vector DNA, and transformed into
chemically-competent E. Coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

ORIGIN

Query Match
Best Local Similarity 100.0%; Score 101; DB 9; Length 495;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTAGAAAATAAACAATAG 60
|
Db 391 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTAGAAAATAAACAATAG 332
|
Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 101
|
```

```
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTAGAAAATAAACAATAG 60
|
Db 392 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTAGAAAATAAACAATAG 333
|
Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 101
|
Db 332 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 292
|

RESULT 11
CC818523/c
LOCUS
DEFINITION
CC818523 496 bp DNA linear GSS 17-JUL-2003
100004L13R Oxytricha plasmid UUGC10 library Sterkiella
histriomuscorum genomic clone UUGC100004L13 R, genomic survey
sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Sterkiella histriomuscorum (Oxytricha trifallax)
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
Stichotrichida; Oxytrichidae; Sterkiella.
REFERENCE
AUTHORS
1 (bases 1 to 496)
Dunn,D., Doak,T., Herrick,G. and Weiss,R.
Paired end reads from plasmid inserts of Oxytricha trifallax
macronuclear chromosomes
Unpublished (2003)
JOURNAL
COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Plate: 0004 row: L column: 13
Seq primer: CACACAGGAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 496.
Location/Qualifiers
1. 496
/organism="Sterkiella histriomuscorum"
/mol_type="genomic DNA"
/db_xref="taxon:94289"
/clone="UUGC100004L13"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Oxytricha plasmid UUGC10 library"
/note="Vector: PWD42nv; Purified macronuclear chromosomal
DNA from Oxytricha trifallax was blunt end-repaired with
T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
oligonucleotides were ligated to the blunt ends in high
molar excess. Vector DNA was prepared from a derivative of
PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
derivative of plasmid R1. The vector was ligated with
adaptors complementary to the insert adaptors and
purified. The sheared, adaptor mouse DNA was annealed to
adaptor vector DNA, and transformed into
chemically-competent E. Coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

ORIGIN

Query Match
Best Local Similarity 100.0%; Score 101; DB 9; Length 496;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTAGAAAATAAACAATAG 60
|
Db 391 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTAGAAAATAAACAATAG 332
|
Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 101
|
```

```

Db      331 GGGTCCGCGCATTTCCTCCCAGAAAGTGCCACCTGCATC 291
|||||
RESULT 12
CcB19854/c
LOCUS   CcB19854       503 bp    DNA     linear     GSS 17-JUL-2003
DEFINITION Oxytricha plasmid UUGC10 library Sterkiella histriomuscorum genomic clone UUGCI0006N08 R, genomic survey sequence.
ACCESION CC819854
VERSION  CcB19854.1 GI:32900533
KEYWORDS
SOURCE   Sterkiella histriomuscorum (Oxytricha trifallax)
ORGANISM Sterkiella histriomuscorum
          Eukaryota; Alveolata; Cilioophora; Spirotrichea; Stichotrichia;
          Stichotrichida; Oxytrichidae; Sterkiella.
REFERENCE Dunn,D., Doak,T., Herrick,G. and Weiss,R.
AUTHORS   Paired end reads from plasmid inserts of Oxytricha trifallax
TITLE     macronuclear chromosomes
JOURNAL   Unpublished (2003)
COMMENT   Contact: Robert B. Weiss
           University of Utah Genome Center
           Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
           84112, USA
           Tel: 801 585 5606
           Fax: 801 585 7177
           Email: ddunn@genetics.utah.edu
           Plate: 0006 row: N column: 08
           Seq primer: CACACAGGAACAAGCATGTACGCC
           Class: plasmid ends
           High quality sequence stop: 503.
FEATURES             Location/Qualifiers
     source            ..503
     /organism="Sterkiella histriomuscorum"
     /mol_type="genomic DNA"
     /db_xref="taxon:94289"
     /clone="UUGC10006N08"
     /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
     /note="Vector: PWD42nv; Purified macronuclear chromosomal DNA from Oxytricha trifallax was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. Vector DNA was prepared from a derivative of pWD42 [gi|4732114|gb|AFI29072.1], a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN
Query Match              100.0%; Score 101; DB 9; Length 503;
Best Local Similarity    100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY   1 AGGGTTATTGTCTCATGCGCGATACATATTTGAATGTATTTAGAATAAATAAACCAATAG 60
|||
DB   410 AGGGTTATTGTCTCATGCGCGATACATATTTGAATGTATTTAGAATAAATAAACCAATAG 351
|||

QY   61 GG GTTC CGCG CAT TTCCCCG AAAAG TGCC ACC TC GC ATC 101
|||
DB   350 GG GTTC CGCG CAT TTCCCCG AAAAG TGCC ACC TC GC ATC 310
|||


ORIGIN
Query Match              100.0%; Score 101; DB 9; Length 503;
Best Local Similarity    100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY   1 AGGGTTATTGTCTCATGCGCGATACATATTTGAATGTATTTAGAATAAATAAACCAATAG 60
|||
DB   410 AGGGTTATTGTCTCATGCGCGATACATATTTGAATGTATTTAGAATAAATAAACCAATAG 351
|||

QY   61 GG GTTC CGCG CAT TTCCCCG AAAAG TGCC ACC TC GC ATC 101
|||
DB   350 GG GTTC CGCG CAT TTCCCCG AAAAG TGCC ACC TC GC ATC 310
|||


RESULT 13
CcB17752/c
LOCUS   CcB17752       518 bp    DNA     linear     GSS 17-JUL-2003
DEFINITION Oxytricha plasmid UUGC10 library Sterkiella histriomuscorum genomic clone UUGCI0002D21 R, genomic survey sequence.
ACCESION CC817128
VERSION  CcB17128.1 GI:32896415
KEYWORDS
SOURCE   Sterkiella histriomuscorum (Oxytricha trifallax)
ORGANISM Sterkiella histriomuscorum
          Eukaryota; Alveolata; Cilioophora; Spirotrichea; Stichotrichia;

```

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Stichotrichida; Oxytrichidae; Sterkiella.
 1 (bases 1 to 518)
 Dunn, D., Doak, T., Herrick, G. and Weiss, R.
 Paired end reads from plasmid inserts of Oxytricha trifallax
 macronuclear chromosomes
 Unpublished (2003)
 Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Plate: 0002 row: D column: 21
 Seq primer: CACACAGGAAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 518.

FEATURES

source
 1..518
 Location/Qualifiers
 /organism="Sterkiella histriomuscorum"
 /mol_type="genomic DNA"
 /db_xref="taxon:94289"
 /clone="UUGC100002D21"
 /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
 /note="Vector: PWD42nv; Purified macronuclear chromosomal
 DNA from Oxytricha trifallax was blunt end-repaired with
 T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
 oligonucleotides were ligated to the blunt ends in high
 molar excess. Vector DNA was prepared from a derivative of
 PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
 derivative of plasmid R1. The vector was ligated with
 adaptors complementary to the insert adaptors and
 purified. The sheared, adapted mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. Coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

ORIGIN

Query Match 100.0%; Score 101; DB 9; Length 518;
 Best Local Similarity 100.0%; Pred. No. 8.4e-19;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTTAGAAAAATAACAATAG 60
 Db 410 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTTAGAAAAATAACAATAG 351

Qy 61 GGGTTCCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 101
 Db 350 GGGTTCCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 310

RESULT 15
 CC817162/c
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

519 bp DNA linear GSS 17-JUL-2003
 100002J19R Oxytricha plasmid UUGC10 library Sterkiella
 histriomuscorum genomic clone UUGC100002J19 R, genomic survey
 sequence.
 CC817162
 CC817162.1 GI:32896449
 GSS.
 Sterkiella histriomuscorum (Oxytricha trifallax)
 Sterkiella histriomuscorum
 Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichida;
 Stichotrichida; Oxytrichidae; Sterkiella.
 1 (bases 1 to 519)
 Dunn, D., Doak, T., Herrick, G. and Weiss, R.
 Paired end reads from plasmid inserts of Oxytricha trifallax
 macronuclear chromosomes
 Unpublished (2003)
 Contact: Robert B. Weiss
 University of Utah Genome Center

University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Plate: 0002 row: J column: 19
 Seq primer: CACACAGGAAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 519.

FEATURES

source
 1..519
 Location/Qualifiers
 /organism="Sterkiella histriomuscorum"
 /mol_type="genomic DNA"
 /db_xref="taxon:94289"
 /clone="UUGC100002J19"
 /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
 /clone_lib="Oxytricha plasmid UUGC10 library"
 /note="Vector: PWD42nv; Purified macronuclear chromosomal
 DNA from Oxytricha trifallax was blunt end-repaired with
 T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
 oligonucleotides were ligated to the blunt ends in high
 molar excess. Vector DNA was prepared from a derivative of
 PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
 derivative of plasmid R1. The vector was ligated with
 adaptors complementary to the insert adaptors and
 purified. The sheared, adapted mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. Coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

ORIGIN

Query Match 100.0%; Score 101; DB 9; Length 519;
 Best Local Similarity 100.0%; Pred. No. 8.4e-19;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTTAGAAAAATAACAATAG 60
 Db 416 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTTAGAAAAATAACAATAG 357

Qy 61 GGGTTCCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 101
 Db 356 GGGTTCCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 316

Search completed: July 14, 2005, 23:22:58
 Job time : 961.667 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:39:07 ; Search time 749.127 Seconds
(without alignments)
6468.225 Million cell updates/sec

Title: US-09-482-682-44_COPY_1_100
Perfect score: 100
Sequence: 1 gacggatcgaggatctcccc.....ctgtccctgtgtgtgtgt 100

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_phi.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100	100.0	3853	6 AR098190	AR098190 Sequence
2	100	100.0	3853	6 AR207832	AR207832 Sequence
3	100	100.0	3853	6 BD009729	BD009729 Tissue sp
4	100	100.0	3986	12 PCDNA3ZEO	X30639 Cloning vec
5	100	100.0	4026	6 AR098191	AR098191 Sequence
6	100	100.0	4026	6 AR207833	AR207833 Sequence
7	100	100.0	4026	6 BD009730	BD009730 Tissue sp
8	100	100.0	4249	6 AR098192	AR098192 Sequence
9	100	100.0	4249	6 AR207834	AR207834 Sequence
10	100	100.0	4249	6 BD009731	BD009731 Tissue sp
11	100	100.0	4341	6 A38214	A38214 Sequence 58
12	100	100.0	4341	6 AX286570	AX286570 Sequence
13	100	100.0	4597	6 AX060344	AX060344 Sequence
14	100	100.0	4840	6 AX133940	AX133940 Sequence
15	100	100.0	5053	6 BD238492	BD238492 Expressio
16	100	100.0	5070	6 AX234391	AX234391 Sequence
17	100	100.0	5070	6 A91754	A91754 Sequence 10
18	100	100.0	5082	6 BD085110	BD085110 Vertebrat
19	100	100.0	5162	6 AX951626	AX951626 Sequence

20	100	100.0	5257	12 CVU89673	U89673 Cloning vec
21	100	100.0	5432	6 BD234590	BD234590 Screening
22	100	100.0	5432	6 AX026821	AX026821 Sequence
23	100	100.0	5446	6 AX319694	AX319694 Sequence
24	100	100.0	5618	6 A44171	A44171 Sequence 1
25	100	100.0	5618	6 AR116416	AR116416 Sequence
26	100	100.0	5618	6 AR222666	AR222666 Sequence
27	100	100.0	5618	6 AR411127	AR411127 Sequence
28	100	100.0	5639	12 AY437643	AY437643 Expressio
29	100	100.0	5651	6 AX211282	AX211282 Sequence
30	100	100.0	5651	6 AX349366	AX349366 Sequence
31	100	100.0	5653	6 I56772	I56772 Sequence 3
32	100	100.0	5653	6 I95540	I95540 Sequence 1
33	100	100.0	5726	12 CVU89672	U89672 Cloning vec
34	100	100.0	5731	6 AX202478	AX202478 Sequence
35	100	100.0	5900	6 AX573107	AX573107 Sequence
36	100	100.0	5995	6 AX685746	AX685746 Sequence
37	100	100.0	6090	6 A63067	A63067 Sequence 11
38	100	100.0	6148	6 BD181637	BD181637 Novel mel
39	100	100.0	6148	6 AX342685	AX342685 Sequence
40	100	100.0	6149	6 BD181638	BD181638 Novel mel
41	100	100.0	6149	6 AX342686	AX342686 Sequence
42	100	100.0	6180	6 AX207724	AX207724 Sequence
43	100	100.0	6186	6 AX211281	AX211281 Sequence
44	100	100.0	6186	6 AX349365	AX349365 Sequence
45	100	100.0	6200	6 BD232461	BD232461 Recombina

ALIGNMENTS

RESULT 1
AR098190
LOCUS AR098190 3853 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 5 from patent US 6074850.
ACCESSION AR098190
VERSION AR098190.1 GI:12807447
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3853)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion polypeptides
JOURNAL Patent: US 6074850-A 5 13-JUN-2000;
FEATURES Location/Qualifiers
source
1..3853
/organism="unknown"
/mol_type="unassigned DNA"

Query Match 100.0%; Score 100; DB 6; Length 3853;
Best Local Similarity 100.0%; Pred. No. 9.4e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GACGGATCGGAGATCTCCCGATCCCGATCGCTCGACTCTCAGTACATCTGCTGTGATG 60
Db	1	GACGGATCGGAGATCTCCCGATCCCGATCGCTCGACTCTCAGTACATCTGCTGTGATG 60
Qy	61	CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTT 100
Db	61	CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTT 100

RESULT 2
AR207832
LOCUS AR207832 3853 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 5 from patent US 6379927.
ACCESSION AR207832
VERSION AR207832.1 GI:21507688
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.


```
Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100
    |||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100

RESULT 6
AR207833
LOCUS 4026 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 19 from patent US 6379927.
ACCESSION AR207833
VERSION AR207833.1 GI:21507689
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4026)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion proteins
JOURNAL Patent: US 6379927-A 19 30-APR-2002;
FEATURES
    Location/Qualifiers
    1..4026
    /organism="unknown"
    /mol_type="unassigned DNA"
ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4026;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
    |||||
Db 1 GACGGATCGGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60

RESULT 7
BD009730
LOCUS 4026 bp DNA linear PAT 31-JAN-2002
DEFINITION Tissue specific expression of retinoblastoma protein.
ACCESSION BD009730
VERSION BD009730.1 GI:18638103
KEYWORDS JP 2001503638-A/4.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 4026)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Tissue specific expression of retinoblastoma protein
JOURNAL Patent: JP 2001503638-A 4 21-MAR-2001;
COMMENT CANI INC
OS Unidentified
PN JP 2001503638-A/4
PD 21-MAR-2001
PF 13-NOV-1997 JP 1998522958
PR 15-NOV-1996 US 08/751517,14-FEB-1997 US 08/801092 PI
DOUGLAS ANTELMAN,RICHARD J GREGORY,KENNETH N WILLS PC
C07H21/04,C07K5/00,A61K38/00,A61K35/12
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1..4026
    /organism='Unidentified'.
    Location/Qualifiers
    1..4026
    /organism="unidentified"
    /mol_type="genomic DNA"
    /db_xref="taxon:32644"
FEATURES
    source
ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4026;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
    |||||
Db 1 GACGGATCGGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100
    |||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100

RESULT 8
AR207833
LOCUS 4249 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 33 from patent US 6074850.
ACCESSION AR207833
VERSION AR207833.1 GI:12807449
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4249)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion polypeptides
JOURNAL Patent: US 6074850-A 33 13-JUN-2000;
FEATURES
    Location/Qualifiers
    1..4249
    /organism="unknown"
    /mol_type="unassigned DNA"
ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4249;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
    |||||
Db 1 GACGGATCGGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100
    |||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100

RESULT 9
AR207834
LOCUS 4249 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 33 from patent US 6379927.
ACCESSION AR207834
VERSION AR207834.1 GI:21507691
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4249)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion proteins
JOURNAL Patent: US 6379927-A 33 30-APR-2002;
FEATURES
    Location/Qualifiers
    1..4249
    /organism="unknown"
    /mol_type="unassigned DNA"
ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4249;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
    |||||
Db 1 GACGGATCGGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100
    |||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100
```


/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="pCDNA3.1/GS vector by Invitrogen Corporation"

ORIGIN

Query Match 100.0%; Score 100; DB 6; Length 4597;
Best Local Similarity 100.0%; Pred. No. 9.2e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
|||||
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
|||||

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100
|||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100
|||||

RESULT 14
AX133940
LOCUS AX133940 4840 bp DNA linear PAT 15-MAY-2001
DEFINITION Sequence 1 from Patent WO0119853.
ACCESSION AX133940
VERSION AX133940.1 GI:14139881
KEYWORDS
SOURCE
ORGANISM
synthetic construct
synthetic construct
other sequences; artificial sequences.
REFERENCE
AUTHORS Hollander,A.P., Barker,M.D. and Kafienah,W.C.
TITLE Cell transfection
JOURNAL Patent: WO 0119853-A 1 22-MAR-2001;
THE UNIVERSITY OF SHEFFIELD (GB)
LOCATION/Qualifiers
FEATURES
source 1..4840
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="This sequence is artificial and is based on well established commercially available vectors that are cited with their vendor withi n the patent applicatio"

ORIGIN

Query Match 100.0%; Score 100; DB 6; Length 4840;
Best Local Similarity 100.0%; Pred. No. 9.2e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
|||||
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
|||||

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100
|||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100
|||||

RESULT 15
BD238492
LOCUS BD238492 5053 bp DNA linear PAT 17-JUL-2003
DEFINITION Expression vectors for stimulating an immune response and methods of using the same.
ACCESSION BD238492
VERSION BD238492.1 GI:33048262
KEYWORDS JP 2002520000-A/18.
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE
AUTHORS Fikes,J.D., Hermanson,G.G., Sette,A., Ishioka,G.Y., Livingston,B. and Chesnut,R.W.
TITLE Expression vectors for stimulating an immune response and methods of using the same
JOURNAL Patent: JP 2002520000-A 18 09-JUL-2002;

EPIMMUNE INC
OS Artificial Sequence
PN JP 2002520000-A/18
PD 09-JUL-2002
PF 13-MAY-1999 JP 2000548449
PR 13-MAY-1998 US 09/078904,15-MAY-1998 US 60/085751 PI
JOHN D FIKES,GARY G HERMANSON,ALESSANDRO
SETTE,GLENN Y ISHIOKA,
PI BRIAN LIVINGSTON,ROBERT W CHESNUT
PC C12N15/09,A61K31/711,A61K39/00,A61K39/12,A61K39/21,A61K39/29,
A61P31/14,
PC A61P31/16,A61P31/20,A61P37/02,C12N15/00
CC vector pEP2
FH Key Location/Qualifiers
FT source 1..5053
/organism="Artificial Sequence".
FEATURES
source 1..5053
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

ORIGIN

Query Match 100.0%; Score 100; DB 6; Length 5053;
Best Local Similarity 100.0%; Pred. No. 9.2e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
|||||
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
|||||

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100
|||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100
|||||

Search completed: July 14, 2005, 14:03:30
Job time : 749.127 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:35:42 ; Search time 140.988 Seconds
(without alignments)
4198.742 Million cell updates/sec

Title: US-09-482-682-44_COPY_1_100
Perfect score: 100
Sequence: 1 gacggatcggagatctccc.....ctgctccctgtgtgtgtt 100

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002as:*
- 7: geneseqn2002bs:*
- 8: geneseqn2003as:*
- 9: geneseqn2003bs:*
- 10: geneseqn2003cs:*
- 11: geneseqn2003ds:*
- 12: geneseqn2004as:*
- 13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	100	100.0	1506	12 ADM41035	Adm41035 Fungus nu
2	100	100.0	1600	2 ADH11349	Adh11349 Vertebrat
3	100	100.0	1782	12 ADM41037	Adm41037 Cytomegal
4	100	100.0	2241	12 ADM41034	Adm41034 Human nuc
5	100	100.0	2294	12 ADM41036	Adm41036 Cytomegal
6	100	100.0	3853	2 AAV40006	Aav40006 Plasmid p
7	100	100.0	4026	2 AAV40007	Aav40007 Plasmid p
8	100	100.0	4249	2 AAV63466	Aav63466 Plasmid p
9	100	100.0	4341	2 AAG62391	Aag62391 Vector pV
10	100	100.0	4341	6 AAS17704	Aas17704 Vector pV
11	100	100.0	4341	6 ABN83143	Abn83143 Plasmid p
12	100	100.0	4597	4 AAF24901	Aaf24901 Nucleotid
13	100	100.0	4639	6 AAD39652	Aad39652 Human sma
14	100	100.0	4840	4 AAF83146	Aaf83146 Complete
15	100	100.0	5015	10 ADB33528	Adb33528 Expressio
16	100	100.0	5053	3 AAZ38633	Aaz38633 pEP2 expr
17	100	100.0	5070	4 AAS12839	Aas12839 DNA sequ
18	100	100.0	5082	2 ADH11417	Adh11417 Plasmid p
19	100	100.0	5162	10 ADF10526	Adf10526 Plasmid p
20	100	100.0	5162	10 ACC44637	Acc44637 Murine rD

21	100	100.0	5172	13 ADS75099	Ads75099 Plasmid p
22	100	100.0	5192	10 ACC44692	Acc44692 Plasmid p
23	100	100.0	5271	10 ABV77540	Abv77540 Plasmid p
24	100	100.0	5283	10 ABV77538	Abv77538 Plasmid p
25	100	100.0	5293	10 ABV77549	Abv77549 Plasmid p
26	100	100.0	5302	12 ADI34681	Adi34681 Nucleotid
27	100	100.0	5304	10 ABV77539	Abv77539 Plasmid p
28	100	100.0	5425	2 ADH11233	Adh11233 Vertebrat
29	100	100.0	5431	6 ABN86685	Abn86685 Nucleotid
30	100	100.0	5431	10 ADE21866	Ade21866 Plasmid v
31	100	100.0	5431	12 ADO05277	Ado05277 pcDNA3 pl
32	100	100.0	5432	3 AAZ89476	Aaz89476 Transgeni
33	100	100.0	5446	6 AAS18619	Aas18619 Renilla l
34	100	100.0	5446	6 ABL53540	AbL53540 Vector pc
35	100	100.0	5446	12 ADN36314	Adn36314 Plasmid p
36	100	100.0	5458	6 ABL58494	AbL58494 Recombina
37	100	100.0	5458	6 ABL58493	AbL58493 Recombina
38	100	100.0	5543	6 ABK88868	Abk88868 Topoisome
39	100	100.0	5543	12 ADE83791	Ade83791 Plasmid p
40	100	100.0	5543	12 ADO06720	Ado06720 Recombina
41	100	100.0	5614	6 ABL58489	AbL58489 Recombina
42	100	100.0	5614	6 ABL58490	AbL58490 Recombina
43	100	100.0	5618	2 AAQ88310	Aaq88310 Plasmid p
44	100	100.0	5651	5 AAI66195	Aai66195 Human FSH
45	100	100.0	5651	6 ABK40237	Abk40237 DNA encod

ALIGNMENTS

RESULT 1

ADMA41035

ID ADM41035 standard; DNA; 1506 BP.

AC ADM41035;

DT 17-JUN-2004 (first entry)

DE Fungus nucleotide sequence SEQ ID NO:3.

XX engrafting foreign replacement cell; implanting foreign replacement cell;
XX growth; differentiation; drug development; vaccine development;
XX tissue transplantation; human disease study; fungus; gene; ds.

OS Unidentified.

PN WO2004027029-A2.

PD 01-APR-2004.

PF 17-SEP-2003; 2003WO-US029251.

PR 19-SEP-2002; 2002US-0411790P.

XX (XIME-) XIMEREX INC.

XX Beschorner WE, Sosa CE, Thompson SC;

XX WPI; 2004-295402/27.

XX Engrafting foreign replacement cells within a fetal non-human mammal,
XX useful in producing chimeric mammals, comprises selectively destroying
XX native cells in a tissue of a fetal non-human mammal host.

XX Disclosure; SEQ ID NO 3; 48pp; English.

XX The present invention describes a method for engrafting foreign
XX replacement cells within a foetal non-human mammal, which comprises
XX selectively destroying native cells in a tissue of a foetal non-human
XX mammal host, where the number of maternal cells of the same tissue is not
XX substantially reduced, and implanting foreign replacement cells in the
XX tissue of the foetal non-human mammal host, where the foreign replacement
XX cells replace destroyed cells of the tissue. The method is useful for

CC facilitating growth and differentiation of foreign cells within a
CC mammalian host, and for producing chimeric mammals that can be used to
CC develop new drugs and vaccine, factors, drugs and tissues for
CC transplantation, also useful to study human diseases. The present
CC sequence represents a nucleotide sequence given in the Sequence Listing
CC of the present invention but not mentioned further within the
CC specification.

XX
SQ Sequence 1506 BP; 454 A; 277 C; 361 G; 414 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 12; Length 1506;
Best Local Similarity 100.0%; Pred. No. 4e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100

RESULT 2
ADH11349
ID ADH11349 standard; DNA; 1600 BP.

XX
AC ADH11349;
XX 11-MAR-2004 (first entry)

DE Vertebrate UNC-53 protein homologue related nucleotide sequence.

XX
KW UNC-53 vertebrate protein homologue; UNC-53; Caenorhabditis elegans;
KW cell shape regulator; cell motility regulator; cell migration;
KW cell behaviour regulator; phenotype; signal transduction pathway;
KW signal transducing protein; signal integrator protein;
KW neuronal regeneration; revascularisation; wound healing;
KW chronic neurodegenerative disease; acute traumatic injury;
KW fibrotic disease; gene; ds.

XX
OS Unidentified.

XX
FN WO9824810-A2.
XX 11-JUN-1998.

XX
PF 03-DEC-1997; 97WO-EP006956.
XX
PR 04-DEC-1996; 96GB-00025283.

XX
PA (JANC) JANSSEN PHARM NV.

XX
PI Plattesuw CJ, Buesa Arjol CM, Deraeymaeker M, Verhasselt P;
PI Pujol NJR, Maertens LJS, Luyten W, Geerts H, Vandekerckhove JS;
PI Geysen J, Bogaert TAOE;

XX
WPI; 1998-362411/31.
DR P-PSDB; ADH11350.

XX
PT Vertebrate homologues of C. elegans UNC-53 protein - useful for, e.g.
PT promoting neuronal regeneration, treating chronic neuro-degenerative
PT diseases or acute traumatic injuries.

XX
PS Disclosure; Page 410-411; 479pp; English.

CC The present invention describes a vertebrate protein homologue of an UNC-
CC 53 protein of Caenorhabditis elegans or a functional equivalent,
CC derivative or precursor of UNC-53. Also described: (1) a cDNA sequence
CC encoding a vertebrate homologue of the C. elegans UNC-53 protein; (2) a
CC nucleic acid which hybridises to the cDNA of (1); (3) vector comprising
CC the cDNA as in (1); (4) a host cell containing the vector as in (3); (5)
CC a transgenic cell, tissue or animal comprising the vector as in (3); (6)

CC a compound identified as an enhancer or inhibitor of the regulation of
CC cell shape, motility, or the direction of cell migration for use as a
CC therapeutic; (7) a method for determination of whether a protein is an
CC inhibitor or enhancer of regulation of cell behaviour, growth, shape or
CC motility or the direction of migration by contacting a host cell
CC expressing a homologue of UNC-53 and determining a change of phenotype;
CC (8) a method for identification of vertebrate homologues of C. elegans
CC unc-53 gene by hybridising a probe of 15-50 bp of the unc-53 sequence to
CC a DNA library; and (9) a method for identification of a protein which is
CC active in the signal transduction pathway of a cell of which a vertebrate
CC homologue of UNC-53 is a component comprising: (i) contacting an extract
CC of a cell with an antibody to the UNC-53 homologue; (ii) identifying an
CC antibody/homologue complex; and (iii) analysing such a complex to
CC identify any non-antibody protein bound to the complex. UNC-53 is a
CC signal transducing or signal integrator protein involved in controlling
CC directionality of cell migration and cell shape in C. elegans. Vertebrate
CC homologues of UNC-53 can be used to promote neuronal regeneration,
CC revascularisation or wound healing, to treat chronic neurodegenerative
CC diseases or acute traumatic injuries or fibrotic diseases. The present
CC sequence is used in the exemplification of the present invention.

XX
SQ Sequence 1600 BP; 397 A; 409 C; 398 G; 396 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 2; Length 1600;
Best Local Similarity 100.0%; Pred. No. 4.1e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100

RESULT 3
ADH41037
ID ADH41037 standard; DNA; 1782 BP.

XX
AC ADH41037;
XX 17-JUN-2004 (first entry)

DE Cytomegalovirus nucleotide sequence SEQ ID NO:5.

XX
KW engrafting foreign replacement cell; implanting foreign replacement cell;
KW growth; differentiation; drug development; vaccine development;
KW tissue transplantation; human disease study; cytomegalovirus; gene; ds.

XX
OS Cytomegalovirus.

XX
FN WO2004027029-A2.
XX 01-APR-2004.

XX
PF 17-SEP-2003; 2003WO-US029251.
XX 19-SEP-2002; 2002US-0411790P.

XX
PA (XIME-) XIMEREX INC.

XX
PI Beschoner WE, Sosa CE, Thompson SC;
XX WPI; 2004-295402/27.

XX
PT Engrafting foreign replacement cells within a fetal non-human mammal,
PT useful in producing chimeric mammals, comprises selectively destroying
PT native cells in a tissue of a fetal non-human mammal host.

XX
PS Disclosure; SEQ ID NO 5; 48pp; English.

XX
CC The present invention describes a method for engrafting foreign

CC replacement cells within a foetal non-human mammal, which comprises
 CC selectively destroying native cells in a tissue of a foetal non-human
 CC mammal host, where the number of maternal cells of the same tissue is not
 CC substantially reduced, and implanting foreign replacement cells in the
 CC tissue of the foetal non-human mammal host, where the foreign replacement
 CC cells replace destroyed cells of the tissue. The method is useful for
 CC facilitating growth and differentiation of foreign cells within a
 CC mammalian host, and for producing chimeric mammals that can be used to
 CC develop new drugs and vaccine, factors, drugs and tissues for
 CC transplantation, also useful to study human diseases. The present
 CC sequence represents a nucleotide sequence given in the Sequence Listing
 CC of the present invention but not mentioned further within the
 CC specification.

XX SQ Sequence 1782 BP; 458 A; 399 C; 451 G; 474 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 12; Length 1782;
 Best Local Similarity 100.0%; Pred. No. 4.2e-26;
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAACTCTGCTCTGATG 60
 Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAACTCTGCTCTGATG 60

Oy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
 Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

RESULT 4

ADM41034
 ID ADM41034 standard; DNA; 2241 BP.

XX AC ADM41034;

DT 17-JUN-2004 (first entry)

XX Human nucleotide sequence SEQ ID NO:2.

XX engrafting foreign replacement cell; implanting foreign replacement cell;
 XX growth; differentiation; drug development; vaccine development;
 XX tissue transplantation; human disease study; human; gene; ds.

XX Homo sapiens.

XX WO2004027029-A2.

XX 01-APR-2004.

XX 17-SEP-2003; 2003WO-US029251.

XX 19-SEP-2002; 2002US-0411790P.

XX (XIME-) XIMEREX INC.

XX Beschornier WE, Sosa CE, Thompson SC;

XX WPI; 2004-295402/27.

XX Engrafting foreign replacement cells within a foetal non-human mammal,
 XX useful in producing chimeric mammals, comprises selectively destroying
 XX native cells in a tissue of a foetal non-human mammal host.

XX Disclosure; SEQ ID NO 2; 48pp; English.

XX The present invention describes a method for engrafting foreign
 CC replacement cells within a foetal non-human mammal, which comprises
 CC selectively destroying native cells in a tissue of a foetal non-human
 CC mammal host, where the number of maternal cells of the same tissue is not
 CC substantially reduced, and implanting foreign replacement cells in the
 CC tissue of the foetal non-human mammal host, where the foreign replacement
 CC cells replace destroyed cells of the tissue. The method is useful for
 CC facilitating growth and differentiation of foreign cells within a

CC mammalian host, and for producing chimeric mammals that can be used to
 CC develop new drugs and vaccine, factors, drugs and tissues for
 CC transplantation, also useful to study human diseases. The present
 CC sequence represents a nucleotide sequence given in the Sequence Listing
 CC of the present invention but not mentioned further within the
 CC specification.

XX SQ Sequence 2241 BP; 498 A; 630 C; 609 G; 504 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 12; Length 2241;
 Best Local Similarity 100.0%; Pred. No. 4.5e-26;
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAACTCTGCTCTGATG 60
 Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAACTCTGCTCTGATG 60

Oy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
 Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

RESULT 5

ADM41036

ID ADM41036 standard; DNA; 2294 BP.

XX AC ADM41036;

DT 17-JUN-2004 (first entry)

XX Cytomegalovirus nucleotide sequence SEQ ID NO:4.

XX engrafting foreign replacement cell; implanting foreign replacement cell;
 XX growth; differentiation; drug development; vaccine development;
 XX tissue transplantation; human disease study; cytomegalovirus; gene; ds.

XX Cytomegalovirus.

XX WO2004027029-A2.

XX 01-APR-2004.

XX 17-SEP-2003; 2003WO-US029251.

XX 19-SEP-2002; 2002US-0411790P.

XX (XIME-) XIMEREX INC.

XX Beschornier WE, Sosa CE, Thompson SC;

XX WPI; 2004-295402/27.

XX Engrafting foreign replacement cells within a foetal non-human mammal,
 XX useful in producing chimeric mammals, comprises selectively destroying
 XX native cells in a tissue of a foetal non-human mammal host.

XX Disclosure; SEQ ID NO 4; 48pp; English.

XX The present invention describes a method for engrafting foreign
 CC replacement cells within a foetal non-human mammal, which comprises
 CC selectively destroying native cells in a tissue of a foetal non-human
 CC mammal host, where the number of maternal cells of the same tissue is not
 CC substantially reduced, and implanting foreign replacement cells in the
 CC tissue of the foetal non-human mammal host, where the foreign replacement
 CC cells replace destroyed cells of the tissue. The method is useful for
 CC facilitating growth and differentiation of foreign cells within a
 CC mammalian host, and for producing chimeric mammals that can be used to
 CC develop new drugs and vaccine, factors, drugs and tissues for
 CC transplantation, also useful to study human diseases. The present
 CC sequence represents a nucleotide sequence given in the Sequence Listing
 CC of the present invention but not mentioned further within the
 CC specification.

SQ Sequence 2294 BP; 466 A; 680 C; 641 G; 507 T; 0 U; 0 Other;
 Query Match 100.0%; Score 100; DB 12; Length 2294;
 Best Local Similarity 100.0%; Pred. No. 4.5e-26;
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GACGGATCGGGAGATCTCCCGATGCCGATATGCTCCCTATGCTCAGTACAATCTGCTCTGATG 60
 DB 1 GACGGATCGGGAGATCTCCCGATGCCGATATGCTCCCTATGCTCAGTACAATCTGCTCTGATG 60
 QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCGCTGCTGTTGTGTT 100
 DB 61 CCGCATAGTTAAGCCAGTATCTGCTCCCGCTGCTGTTGTGTT 100
 RESULT 6
 AAV40006
 ID AAV40006 standard; DNA; 3853 BP.
 XX AC AAV40006;
 XX 27-AUG-2003 (revised)
 DT 15-FEB-1999 (first entry)
 XX Plasmid pCTM.
 XX E2F; transcription factor; human; retinoblastoma protein RB;
 KW bladder cancer; restenosis; angioplasty; diabetic retinopathy;
 KW thyroid hyperplasia; Grave's disease; psoriasis;
 KW benign prostatic hypertrophy; Li-Fraumeni syndrome;
 KW peripheral vascular disease; therapy; plasmid pCTM; ss.
 XX Human cytomegalovirus.
 OS mastadenovirus.
 OS unidentified bacteriophage; T7.
 OS Macaca mulatta; polyoma virus.
 OS Bos taurus.
 OS Chimeric.
 XX Key Location/Qualifiers
 FT promoter 209..864
 FT /tag= a
 FT /note= "CMV promoter"
 FT misc_feature 907..1131
 FT /tag= b
 FT /function= "tripartite leader sequence"
 FT promoter 1132..1149
 FT /tag= c
 FT /note= "SP6 promoter"
 FT misc_feature 1679..3853
 FT /tag= d
 FT /note= "pUC19 backbone H3 to AatII"
 FT CDS complement (2857..3717)
 FT /tag= e
 FT /note= "AMP-ORF"
 XX WO9821228-A1.
 XX 22-MAY-1998.
 XX 13-NOV-1997; 97WO-US021821.
 XX 15-NOV-1996; 96US-00751517.
 XX 14-FEB-1997; 97US-00801092.
 XX (CANJ-) CANJI INC.
 XX Antelman D, Gregory RJ, Wills KN;
 XX WPI; 1998-297858/26.
 XX New fusion polypeptide of, e.g. transcription factor - used to treat,

e.g. hyper-proliferative disease such as cancer and restenosis.
 Example 1; Fig 4; 91pp; English.
 This is the nucleotide sequence of pCTM, a plasmid which contains a CMV promoter, a tripartite adenovirus leader flanked by T7 and SP6 promoters, and a multiple cloning site with a bovine growth hormone polyA site and downstream SV40 polyA site. It has been used as a vector for the expression of fusion proteins of the invention that comprise retinoblastoma protein (BP, see AAW62465) and E2F transcription factor (see AAW62464). Such fusion proteins, particularly expressed from gene therapy vectors, are used to treat hyperproliferative conditions, specifically cancer (particularly of the bladder) or restenosis. They are more effective in repressing transcription of the E2F promoter than RB alone and cause cell-cycle arrest in a variety of cells. (Updated on 27-AUG-2003 to correct OS field.)
 SQ Sequence 3853 BP; 936 A; 986 C; 942 G; 989 T; 0 U; 0 Other;
 Query Match 100.0%; Score 100; DB 2; Length 3853;
 Best Local Similarity 100.0%; Pred. No. 5.2e-26;
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GACGGATCGGGAGATCTCCCGATGCCGATATGCTCCCTATGCTCAGTACAATCTGCTCTGATG 60
 DB 1 GACGGATCGGGAGATCTCCCGATGCCGATATGCTCCCTATGCTCAGTACAATCTGCTCTGATG 60
 QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCGCTGCTGTTGTGTT 100
 DB 61 CCGCATAGTTAAGCCAGTATCTGCTCCCGCTGCTGTTGTGTT 100
 RESULT 7
 AAV40007
 ID AAV40007 standard; DNA; 4026 BP.
 XX AC AAV40007;
 XX 27-AUG-2003 (revised)
 DT 15-FEB-1999 (first entry)
 XX Plasmid pCTMI.
 XX E2F; transcription factor; human; retinoblastoma protein RB;
 KW bladder cancer; restenosis; angioplasty; diabetic retinopathy;
 KW thyroid hyperplasia; Grave's disease; psoriasis;
 KW benign prostatic hypertrophy; Li-Fraumeni syndrome;
 KW peripheral vascular disease; therapy; plasmid pCTMI; ss.
 XX Human cytomegalovirus.
 OS mastadenovirus.
 OS unidentified bacteriophage; T7.
 OS Macaca mulatta; polyoma virus.
 OS Bos taurus.
 OS Chimeric.
 XX Key Location/Qualifiers
 FT promoter 209..864
 FT /tag= a
 FT /note= "CMV promoter"
 FT misc_feature 907..1074
 FT /tag= b
 FT /function= "tripartite leader sequence"
 FT intron 1075..1253
 FT /tag= c
 FT /note= "hybrid SV40 late intron"
 FT promoter 1305..1322
 FT /tag= d
 FT /note= "SP6 promoter"
 FT misc_feature 1851..4026
 FT /tag= e
 FT /note= "pUC19 backbone H3 to AatII"

```

FT CDS complement(3032. .3890)
FT /*tag= f
FT /*note= "AMP-ORF"
XX
XX WO9821228-A1.
XX
XX 22-MAY-1998.
XX
XX 13-NOV-1997; 97WO-US021821.
XX
XX 15-NOV-1996; 96US-00751517.
XX 14-FEB-1997; 97US-00801092.
XX (CANJ-) CANJI INC.
XX
XX Antelman D, Gregory RJ, Wills KN;
XX
XX WPI; 1998-297858/26..
XX
XX New fusion polypeptide of, e.g. transcription factor - used to treat,
XX e.g. hyper-proliferative disease such as cancer and restenosis.
XX
XX Example 1; Fig 6; 91pp; English.
XX
XX This is the nucleotide sequence of pCTMI, a plasmid that was constructed
XX from pCTM (see AAV40006) by digesting pCTM with XhoI and NotI and
XX subcloning a 180 bp intron XhoI-NotI fragment from a pCMV-beta-gal
XX vector. Plasmid pCTMI has been used as a vector for the expression of
XX fusion proteins of the invention that comprise retinoblastoma protein
XX (BP, see AAW62465) and E2F transcription factor (see AAW62464). Such
XX fusion proteins, particularly expressed from gene therapy vectors, are
XX used to treat hyperproliferative conditions, specifically cancer
XX (particularly of the bladder) or restenosis. They are more effective in
XX repressing transcription of the E2F promoter than RB alone and cause cell
XX -cycle arrest in a variety of cells. (Updated on 27-AUG-2003 to correct
XX OS field.)
XX
XX SQ Sequence 4026 BP; 978 A; 1021 C; 982 G; 1045 T; 0 U; 0 Other;
XX
XX Query'Match 100.0%; Score 100; DB 2; Length 4026;
XX Best Local Similarity 100.0%; Pred. No. 5.3e-26;
XX Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Oy 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
XX |||||
XX Db 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
XX |||||
XX Oy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100
XX |||||
XX Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100
XX |||||
XX
XX RESULT 8
XX AAV63466
XX ID AAV63466 standard; DNA; 4249 BP.
XX
XX AC AAV63466;
XX
XX 27-AUG-2003 (revised)
XX 15-FEB-1999 (first entry)
XX
XX DE Plasmid pCTMIE.
XX
XX E2F; transcription factor; human; retinoblastoma protein RB;
XX bladder cancer; restenosis; angioplasty; diabetic retinopathy;
XX thyroid hyperplasia; Grave's disease; psoriasis;
XX benign prostatic hypertrophy; Li-Fraumeni syndrome;
XX peripheral vascular disease; therapy; plasmid pCTMIE; ss.
XX
XX OS Human cytomegalovirus.
XX OS mastadenovirus.
XX OS unidentified bacteriophage; T7.
XX OS unidentified bacteriophage; SP6.

```

```

OS OS Macaca mulatta; polyoma virus.
OS OS Bos taurus.
OS OS Chimeric.
XX
XX Key Location/Qualifiers
XX FT Promoter 209..864
XX FT /*tag= a
XX FT /*note= "CMV promoter"
XX FT misc_feature 907..1074
XX FT /*tag= b
XX FT /*function= "tripartite leader sequence"
XX FT intron 1081..1145
XX FT /*tag= c
XX FT /*note= "hybrid SV40 late intron"
XX FT mRNA 1164..1366
XX FT /*tag= d
XX FT /*note= "early mRNA"
XX FT enhancer 1261..1332
XX FT /*tag= e
XX FT /*note= "72 bp tandem repeat enhancer"
XX FT enhancer 1333..1404
XX FT /*tag= f
XX FT /*note= "72 bp tandem repeat enhancer"
XX FT misc_binding 1366
XX FT /*tag= g
XX FT /*note= "T antigen binding site"
XX FT intron 1372..1478
XX FT /*tag= h
XX FT /*note= "hybrid SV40 late intron"
XX FT promoter 1530..1545
XX FT /*tag= i
XX FT /*note= "SP6 promoter"
XX FT misc_feature 2075..4249
XX FT /*tag= j
XX FT /*note= "pUC19 backbone H3 to AatII"
XX FT CDS complement(3255..4113)
XX FT /*tag= k
XX FT /*note= "AMP-ORF"
XX
XX WO9821228-A1.
XX
XX 22-MAY-1998.
XX
XX 13-NOV-1997; 97WO-US021821.
XX
XX 15-NOV-1996; 96US-00751517.
XX 14-FEB-1997; 97US-00801092.
XX (CANJ-) CANJI INC.
XX
XX Antelman D, Gregory RJ, Wills KN;
XX
XX WPI; 1998-297858/26..
XX
XX New fusion polypeptide of, e.g. transcription factor - used to treat,
XX e.g. hyper-proliferative disease such as cancer and restenosis.
XX
XX Example 1; Fig 8; 91pp; English.
XX
XX This is the nucleotide sequence of pCTMIE, a plasmid that was constructed
XX by amplifying the SV40 enhancer from SV40 viral DNA by PCR, digesting the
XX amplified product with BglII and inserting into BamHI-digested plasmid
XX pCTMI (see AAV40007). Plasmid pCTMIE has been used as a vector for the
XX expression of fusion proteins of the invention that comprise
XX retinoblastoma protein (BP, see AAW62465) and E2F transcription factor
XX (see AAW62464). Such fusion proteins, particularly expressed from gene
XX therapy vectors, are used to treat hyperproliferative conditions,
XX specifically cancer (particularly of the bladder) or restenosis. They are
XX more effective in repressing transcription of the E2F promoter than RB
XX alone and cause cell-cycle arrest in a variety of cells. (Updated on 27-
XX AUG-2003 to correct OS field.)
XX
XX SQ Sequence 4249 BP; 1020 A; 1074 C; 1048 G; 1107 T; 0 U; 0 Other;

```

Query Match 100.0%; Score 100; DB 2; Length 4249;
Best Local Similarity 100.0%; Pred. No. 5.4e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTCATG 60
DB 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTCATG 60

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100
DB 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100

RESULT 9
AAQ62391
ID AAQ62391 standard; DNA; 4341 BP.

XX AC AAQ62391;
XX
DT 25-MAR-2003 (revised)
DT 18-NOV-1994 (first entry)
XX
XX
DE Vector pVAC1.
XX
KW Vector; pVAC1; pRC/RSV; leader sequence; termination signal;
KW fusion protein; pSfi/NotI; pSfi/NotI; pSfi/NotI; pSfi/NotI; pSfi/NotI;
KW single chain; Fv; murine antibody; retroviral; envelope; plasmid;
KW vaccine; ss.
XX
XX
OS Synthetic.

XX
FH Key Location/Qualifiers
FT misc_RNA complement (1. .775)
FT /*tag= c
FT /note= "Claim 9"
FT misc_RNA 606. .780
FT /*tag= b
FT /note= "Claim 8"
FT misc_RNA 606. .716
FT /*tag= a
FT /note= "Claim 7"

XX
XX WO9408008-A1.
XX
XX 14-APR-1994.
XX
XX 04-OCT-1993; 93WO-GB002054.
XX
XX 02-OCT-1992; 92GB-00020808.
XX
XX (MEDI-) MEDICAL RES COUNCIL.
XX
XX Hawkins RE, Russell SJ, Stevenson FK, Winter GP;
XX
XX WPI; 1994-135575/16.
XX
XX
XX Modulating immune response to a disease marker - by administering a
XX vector which expresses the disease marker to interact with the immune
XX system.
XX
XX Claim 10; Fig 7; 77pp; English.

XX
XX This sequence represents the vector pVAC1. This vector is based on the
XX commercially available vector pRC/RSV. Leader sequences and termination
XX signals were introduced into the vector to allow for production of fusion
XX proteins. The vector, pSfi/NotI, was modified to replace the pSfi/NotI
XX leader with the human immunoglobulin VH1 leader sequence that permits the
XX encoding of an SfiI cloning site without modification of the amino acid
XX sequence. This fragment was then cloned as an EcoRI/Bln-HindIII
XX fragment into NotI/Bln-HindIII cut vector pRC/RSV to give pVAC1. The
XX single chain Fv for an individual patient can be inserted within the VH1
XX leader sequence. This plasmid when encoding a single chain murine

CC antibody/retroviral envelope fusion protein can be used as a plasmid
CC vaccine and it induces a strong humoral response to the antibody moiety
CC in BALB/c mice. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 4341 BP; 1032 A; 1099 C; 1091 G; 1119 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 2; Length 4341;
Best Local Similarity 100.0%; Pred. No. 5.4e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTCATG 60
DB 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTCATG 60

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100
DB 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100

RESULT 10
AAS17704
ID AAS17704 standard; DNA; 4341 BP.

XX AC AAS17704;
XX
DT 12-MAR-2002 (first entry)
XX
XX Vector pVAC1 encoding a DNA vaccine.
XX
XX Cytostatic; vaccine; tetanus toxin; Frc; tumour; CTL; PCR primer; pVAC1;
XX ds.
XX
XX Clostridium tetani.
XX Homo sapiens.
XX Synthetic.
XX Cauliflower mosaic virus.
XX
XX WO200179510-A1.
XX
XX 25-OCT-2001.
XX
XX 17-APR-2001; 2001WO-GB001719.
XX
XX 17-APR-2000; 2000GB-00009470.
XX
XX (CANC-) CANCER RES VENTURES LTD.
XX
XX Rice J, Stevenson F;
XX
XX WPI; 2002-086370/09.
XX
XX
XX Nucleic acid construct, useful to immunize against various diseases
XX including cancer, expresses the first domain of tetanus toxin Frc fused
XX to a disease peptide antigen to provide a vaccine.
XX
XX Disclosure; Fig 4; 71pp; English.

XX
XX The invention relates to a nucleic acid construct for delivery into
XX living cells in vivo, to induce an immune response to a disease peptide
XX antigen, where the construct directs expression of a fusion protein
XX comprising the peptide antigen and the first domain of Frc. Also included
XX are a nucleic acid vector comprising the above construct, a host cell
XX comprising the above construct or vector and a method of producing a
XX nucleic acid construct for inducing an immune response. The method
XX comprises identifying a nucleic acid sequence encoding a disease peptide
XX antigen comprising epitopes characteristic of the disease, cloning the
XX nucleic acid sequence, introducing the cloned nucleic acid into a vector
XX which allows the antigen to be expressed as a fusion with a first domain
XX Frc from tetanus toxin, and optionally isolating the construct from the
XX vector. The construct or vector is used as a vaccine to induce an immune
XX response, particularly to tumour antigens. The present sequence is vector
XX pVAC1 which encodes a vaccine of the invention

SQ Sequence 4341 BP; 1033 A; 1099 C; 1090 G; 1119 T; 0 U; 0 Other;
Query Match 100.0%; Score 100; DB 6; Length 4341;
Best Local Similarity 100.0%; Pred. No. 5.4e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCGATCCCTATGCTCGACTCTCAGTACAAATCTGCTCTGATG 60
DB 1 GACGGATCGGAGATCTCCGATCCCTATGCTCGACTCTCAGTACAAATCTGCTCTGATG 60

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGCTGTTGTT 100
DB 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGCTGTTGTT 100

RESULT 11
ABN83143
ID ABN83143 standard; DNA; 4341 BP.
XX
AC ABN83143;
XX
DT 10-SEP-2002 (first entry)
XX
DE Plasmid pVAC1 complete sequence.
XX
KW Immune response; plant viral coat protein; pVAC1; cytostatic; virucide;
KW cancer; B cell malignancy; ds.
XX
OS Synthetic.
XX
PN WO200240513-A2.
XX
PD 23-MAY-2002.
XX
PF 20-NOV-2001; 2001WO-GB005142.
XX
PR 20-NOV-2000; 2000GB-00028319.
XX
PA (CANC-) CANCER RES VENTURES LTD.
XX
PI Savelyeva N, Stevenson F;
XX
DR WPI; 2002-500202/53.
XX
PS Nucleic acid construct for delivery into living cells as a vaccine,
PT useful for treating e.g. cancer, directs the expression of a fusion
PT protein comprising an antigen and an adjuvant sequence derived from a
PT plant viral coat protein.
XX
XX
XX Example 3; Fig 7; 84pp; English.
XX
CC The invention relates to a novel nucleic acid construct for inducing an
CC immune response in vivo to an antigen, capable of directing the
CC expression of a fusion protein that comprises an antigen and an adjuvant
CC sequence derived from a plant viral coat protein. The construct of the
CC invention has cytostatic and virucide activity. The nucleic acid
CC construct is useful for inducing an immune response in a patient, for
CC vaccinating a patient against an infectious disease caused by an antigen
CC derived from a pathogen e.g. a virus, for treating a cancer patient or a
CC patient with a predisposition to cancer and for treating a patient having
CC a B cell malignancy, where the construct is encapsulated, and optionally,
CC a second nucleic acid sequence encoding a further immunomodulatory
CC polypeptide is administered to the patient. The construct is also useful
CC in medical treatment, and in the preparation of a vaccine for treating or
CC preventing a disease state associated with the antigen. The sequence
CC shows the complete sequence of vector pVAC1
XX
SQ Sequence 4341 BP; 1033 A; 1099 C; 1090 G; 1119 T; 0 U; 0 Other;
Query Match 100.0%; Score 100; DB 6; Length 4341;
Best Local Similarity 100.0%; Pred. No. 5.4e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCGATCCCTATGCTCGACTCTCAGTACAAATCTGCTCTGATG 60
DB 1 GACGGATCGGAGATCTCCGATCCCTATGCTCGACTCTCAGTACAAATCTGCTCTGATG 60

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGCTGTTGTT 100
DB 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGCTGTTGTT 100

RESULT 12
AAF24901
ID AAF24901 standard; DNA; 4597 BP.
XX
AC AAF24901;
XX
DT 20-APR-2001 (first entry)
XX
DE Nucleotide sequence of the plasmid pCDNA3.1/GS.
XX
KW Microsphere; dihydrazide; hyaluronic acid; inflammatory response;
KW myocardial ischemia; cardiac angiogenesis; haemophilia;
KW vascular endothelial growth factor; VEGF; ss.
XX
OS Synthetic.
XX
PN WO200078358-A2.
XX
PD 28-DEC-2000.
XX
PF 19-JUN-2000; 2000WO-US016837.
XX
PR 18-JUN-1999; 99US-0140260P.
XX
PA (COLL-) COLLABORATIVE GROUP LTD.
XX
PI Chen W;
XX
DR WPI; 2001-071363/08.
XX
PT Hyaluronic acid micro spheres for use in gene therapy of myocardial
PT ischemia and hemophilia, comprising dihydrazide derivatized hyaluronic
PT acids crosslinked to nucleic acids.
XX
PS Example 1; Page 36-38; 38pp; English.
XX
CC The specification describes a microsphere comprising dihydrazide
CC derivatized hyaluronic acid crosslinked to a nucleic acid (NA). The
CC microspheres cause reduced inflammatory responses, and have increased
CC safety and biodegradability. The microspheres are useful for transfecting
CC a cell of a subject and for treating a subject having myocardial
CC ischemia, by increasing cardiac angiogenesis. They are also useful for
CC treating haemophilia. The present sequence represents the plasmid
CC pCDNA3.1/GS, into which is inserted a polynucleotide sequence which is
CC crosslinked to hyaluronic acid. The polynucleotide sequence encodes a
CC vascular endothelial growth factor (VEGF)
XX
SQ Sequence 4597 BP; 1062 A; 1214 C; 1206 G; 1115 T; 0 U; 0 Other;
Query Match 100.0%; Score 100; DB 4; Length 4597;
Best Local Similarity 100.0%; Pred. No. 5.5e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCGATCCCTATGCTCGACTCTCAGTACAAATCTGCTCTGATG 60
DB 1 GACGGATCGGAGATCTCCGATCCCTATGCTCGACTCTCAGTACAAATCTGCTCTGATG 60

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGCTGTTGTT 100
DB 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGCTGTTGTT 100

RESULT 13
AAD39652

XX 27-FEB-2002; 2002US-0360274P.
XX (MERI) MERCK & CO INC.
XX Espeseth AS, Ferrer M, Flores OA, Hazuda DJ, Inglesse J;
XX Miller MD, Register B, Shi X, Simon AJ, Zuck PD;
XX WPI; 2003-689968/65.
XX DNA encoding a fusion protein of amyloid precursor protein, useful in
XX screening for anti-Alzheimer agents, comprises a fused transcription
XX factor.
XX Disclosure; Fig 32B-F; 193pp; English.
XX The present invention describes a DNA molecule (I) that encodes a fusion
XX protein (FP) comprising: (i) an amino acid sequence of amyloid precursor
XX protein (APP), either the wild type, Swedish or NFEV versions; and (ii) a
XX transcription factor (TF), fused in frame to the C-terminus of (i). Also
XX described: (1) an expression vector containing (I); (2) a eukaryotic cell
XX containing (I); and (3) methods for identifying a compound (A) that
XX inhibits processing of APP, using the cells of (2). (I) has neurotropic and
XX neuroprotective activities. (I) can be used to produce eukaryotic cells
XX that express FP and are useful in screening for agents that inhibit
XX processing of APP. The agents are potentially useful for the treatment or
XX prevention of Alzheimer's disease. Cells that express FP can screen for
XX inhibitors of: (a) beta- and gamma-secretases; and (b)
XX cytoplasmic/extracellular APP signaling in a single assay. Cell-based
XX assays may be free of interference from alpha-secretase activity and are
XX homogeneous (no chromatography, immunoprecipitation or washing required)
XX so well suited to high-throughput screening. The present sequence
XX represents a plasmid nucleotide sequence from the present invention.
SQ Sequence 5015 BP; 1167 A; 1297 C; 1279 G; 1272 T; 0 U; 0 Other;
Query Match 100.0%; Score 100; DB 10; Length 5015;
Best Local Similarity 100.0%; Pred. No. 5.6e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GACGGATCGGAGATCTCCGATCCCTATGGTCGACTTCAGTACATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCGATCCCTATGGTCGACTTCAGTACATCTGCTCTGATG 60
Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCTGCTTGTTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCTGCTTGTTGTT 100

Search completed: July 14, 2005, 07:01:38
Job time : 141.038 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 05:15:57 ; Search time 952.146 Seconds
(without alignments)
3997.736 Million cell updates/sec

Title: US-09-482-682-44_COPY_1_100

Perfect score: 100

Sequence: 1 gacggatcgaggatctccc.....ctgctccctgtgtgtgtt 100

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST:*

1: gb_est1.*
2: gb_est2.*
3: gb_hcc.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gss1.*
9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	60.0	602	8	B67169 CP00047A Cp
2	55.6	55.6	694	8	B2052929 jnr13g03.
3	55.6	55.6	696	8	B2050328 jnr42c12.
4	55.6	55.6	717	8	B2054067 jnr38b09.
5	53.6	53.6	348	2	AW409112 sal10h5 S
6	53.4	53.4	343	1	AL715724 AL715724
7	53.4	53.4	345	1	AL714571 AL714571
8	53.4	53.4	761	7	CK119397 212c09.p1
9	53.4	53.4	766	7	CK120360 207j04.p1
10	53.4	53.4	788	7	CK117844 209p08.p1
11	53.4	53.4	898	9	CL141237 ISB1-118J
12	53.4	53.4	899	9	CL140877 ISB1-118B
13	53.4	53.4	1009	9	CL123953 ISB1-84J1
14	53.2	53.2	814	8	AQ914559 nbe0049M
15	53	53.0	675	8	B2051815 jnr57d03.
16	53	53.0	679	8	B2052857 jnr13g03.
17	53	53.0	700	8	B2050646 jnr66f08.
18	53	53.0	701	8	B2052015 jnr56b03.
19	53	53.0	708	8	B2054793 jnr33g03.
20	53	53.0	709	8	B2053587 jnr98d01.
21	53	53.0	712	8	B2054005 jnr38b09.
22	52.8	52.8	451	8	AQ863966 nbe0022E
23	52.6	52.6	399	8	AQ075099 CIT-HSP-2
24	52.4	52.4	700	8	B2049113 jnr21d02.

25 52.4 52.4 708 8 BZ050047
26 51.6 51.6 328 9 CC819886
27 51.6 51.6 351 9 CC818492
28 51.6 51.6 358 9 CC817661
29 51.6 51.6 364 9 CC817805
30 51.6 51.6 364 9 CC818511
31 51.6 51.6 364 9 CC818574
32 51.6 51.6 364 9 CC819049
33 51.6 51.6 369 9 CC817069
34 51.6 51.6 374 9 CC817074
35 51.6 51.6 374 9 CC820036
36 51.6 51.6 395 9 CC817652
37 51.6 51.6 403 9 CC817682
38 51.6 51.6 403 9 CC817837
39 51.6 51.6 414 9 CC819240
40 51.6 51.6 419 9 CC818384
41 51.6 51.6 420 9 CC817834
42 51.6 51.6 426 9 CC817720
43 51.6 51.6 437 9 CC819820
44 51.6 51.6 441 9 CC818421
45 51.6 51.6 443 9 CC817769

ALIGNMENTS

RESULT 1
B67169
LOCUS B67169 602 bp DNA linear GSS 12-MAY-2000
DEFINITION CP00047A CpIOWAgDNA2 Cryptosporidium parvum genomic, genomic survey sequence.
ACCESSION B67169
VERSION B67169.1 GI:2642750
KEYWORDS GSS.
SOURCE Cryptosporidium parvum
ORGANISM Cryptosporidium parvum
REFERENCE 1 (bases 1 to 602)
AUTHORS Strong, W.B. and Nelson, R.G.
TITLE Preliminary profile of the Cryptosporidium parvum genome: an expressed sequence tag and genome survey sequence analysis
JOURNAL Mol. Biochem. Parasitol. 107 (1), 1-32 (2000)
MEDLINE 20183851
PUBMED 10717299
COMMENT Contact: Nelson, R. G.
Depts. of Medicine & Pharmaceutical Chemistry
San Francisco General Hospital-University of California, San Francisco
Box 0811, San Francisco, CA 94143-0811, USA
Tel: 415 206 8846
Fax: 415 206 3353
Email: malariad@ucsf.edu
Submitted sequence has been edited to remove vector sequences 5' to the insert, to correct miscalled bases and assign uncalled (N) bases throughout the sequence, and to terminate when base-calling became ambiguous.
Seq primer: T7
Class: shotgun
High quality sequence stop: 602.
Location/Qualifiers
1. .602
/organism="Cryptosporidium parvum"
/mol_type="genomic DNA"
/strain="IOWA"
/db_xref="taxon:5807"
/lab_host="E. coli XL2 Blue MRF"
/clone_lib="CpIOWAgDNA2"
/notes="Vector: pCR-Script Amp SK+; Site_1: SrfI; C. parvum (IOWA isolate) genomic DNA was hydrodynamically sheared to produce fragments having a tight size distribution between 2-4 kb by Dr. Yvonne Thorstenson of the Stanford DNA Sequencing and Technology Center

(http://sequence-www.stanford.edu/group/techdev/shear.htm)

The randomly sheared gDNA was chromatographed on Sephacryl S-400 to remove any small fragments and DNA eluting in the void volume was blunt-ended with T4 DNA Polymerase, treated with alkaline phosphatase to prevent the ligation of multiple fragments, ligated to Srf I-digested PCR-Script Amp (SK+) vector and transformed into E. coli strain XL10 Gold. Recombinant clones from the first plating of the library were selected for sequence analysis using T3 and T7 primers."

ORIGIN

Query Match .60.0%; Score 60; DB 8; Length 602;

Best Local Similarity 100.0%; Pred. No. 2.4e-10;

Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 41 CAGTACAACTGCTCTGATGCGCATAGTTAAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

Db 1 CAGTACAACTGCTCTGATGCGCATAGTTAAAGCCAGTATCTGCTCCCTGCTTGTGTT 60

RESULT 2

BZ052929/c

LOCUS

DEFINITION jnr13g03.g1 B.oleracea001 Brassica oleracea genomic, genomic survey

ACCSSION BZ052929

VERSION BZ052929.1 GI:23654922

KEYWORDS GSS.

SOURCE Brassica oleracea

ORGANISM Brassica oleracea

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

1 (bases 1 to 694)

Delehaunty,K., Fewell,G., Fulton,L., McCombie,W.R., Miner,T.,

Nash,W., Rabinowicz,P.D. and Wilson,R.K.

Whole genome shotgun reads from Brassica oleracea

Unpublished (2002)

Contact: Richard K. Wilson

Genome Sequencing Center

Washington University School of Medicine

Email: submissions@watson.wustl.edu

Plate: jnr13 row: 9 column: 03

Seq primer: -28RppOT reverse

Class: shotgun

High quality sequence start: 32

High quality sequence stop: 551.

FEATURES

source

1..694

/organism="Brassica oleracea"

/mol_type="genomic DNA"

/db_xref="taxon:3712"

/clone_lib="B.oleracea001"

/note="Vector: pOTw13; Whole genome shotgun library from flowering buds. DNA was purified from a crude nuclear prep using Brassica oleracea T01000DH3 buds provided by Thomas Osborn at the University of Wisconsin. Genomic DNA was provided by Pablo Rabinowicz (CSHL) and the shotgun library prepared at Washington University Genome Sequencing Center."

ORIGIN

Query Match 55.6%; Score 55.6; DB 8; Length 694;

Best Local Similarity 77.9%; Pred. No. 9e-09;

Matches 67; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 3 CGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAACTGCTCTGATGCC 62

Db 324 CGGATCGATAGGTCCTGGAGCTAGTTATGGTGCATCTCAGTACAACTGCTCTGATGCC 265

Qy 63 GCATAGTTAAGCCAGTATCTGCTCCC 88

Db 264 GCATAGTTAAGCCAGCCCGACACCC 239

RESULT 3

BZ050328

LOCUS

DEFINITION jnr42c12.b1 B.oleracea001 Brassica oleracea genomic, genomic survey

ACCSSION BZ050328

VERSION BZ050328.1 GI:23649718

KEYWORDS GSS.

SOURCE Brassica oleracea

ORGANISM Brassica oleracea

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

1 (bases 1 to 696)

Delehaunty,K., Fewell,G., Fulton,L., McCombie,W.R., Miner,T.,

Nash,W., Rabinowicz,P.D. and Wilson,R.K.

Whole genome shotgun reads from Brassica oleracea

Unpublished (2002)

Contact: Richard K. Wilson

Genome Sequencing Center

Washington University School of Medicine

Email: submissions@watson.wustl.edu

Plate: jnr42 row: C column: 12

Seq primer: -21UPpOT forward

Class: shotgun

High quality sequence start: 35

High quality sequence stop: 180.

FEATURES

source

1..696

/organism="Brassica oleracea"

/mol_type="genomic DNA"

/db_xref="taxon:3712"

/clone_lib="B.oleracea001"

/note="Vector: pOTw13; Whole genome shotgun library from flowering buds. DNA was purified from a crude nuclear prep using Brassica oleracea T01000DH3 buds provided by Thomas Osborn at the University of Wisconsin. Genomic DNA was provided by Pablo Rabinowicz (CSHL) and the shotgun library prepared at Washington University Genome Sequencing Center."

ORIGIN

Query Match 55.6%; Score 55.6; DB 8; Length 696;

Best Local Similarity 77.9%; Pred. No. 9e-09;

Matches 67; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 3 CGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAACTGCTCTGATGCC 62

Db 45 CGGATCGATAGGTCCTGGAGCTAGTTATGGTGCATCTCAGTACAACTGCTCTGATGCC 104

Qy 63 GCATAGTTAAGCCAGTATCTGCTCCC 88

Db 105 GCATAGTTAAGCCAGCCCGACACCC 130

Qy 63 GCATAGTTAAGCCAGTATCTGCTCCC 88

Db 105 GCATAGTTAAGCCAGCCCGACACCC 130

Qy 63 GCATAGTTAAGCCAGTATCTGCTCCC 88

Db 105 GCATAGTTAAGCCAGCCCGACACCC 130

RESULT 4

BZ054067/c

LOCUS

DEFINITION jnr38b09.g1 B.oleracea001 Brassica oleracea genomic, genomic survey

ACCSSION BZ054067

VERSION BZ054067.1 GI:23657216

KEYWORDS GSS.

SOURCE Brassica oleracea

ORGANISM Brassica oleracea

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

1 (bases 1 to 717)

Delehaunty,K., Fewell,G., Fulton,L., McCombie,W.R., Miner,T.,

TITLE Nash, W., Rabinowicz, P.D. and Wilson, R.K.
JOURNAL Whole genome shotgun reads from *Brassica oleracea*
COMMENT Unpublished (2002)
Contact: Richard K. Wilson

Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Plate: jnr38 row: b column: 09
Seq primer: -28RPOT reverse
Class: shotgun
High quality sequence start: 87
High quality sequence stop: 543.

FEATURES

source

1. .717
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/db_xref="taxon:3712"
/clone_lib="B.oleracea001"
/note="Vector: pOTw13; Whole genome shotgun library from
flowering buds. DNA was purified from a crude nuclear
prep using *Brassica oleracea* T01000DH3 buds provided by
Thomas Osborn at the University of Wisconsin. Genomic
DNA was provided by Pablo Rabinowicz (CSHL) and the
shotgun library prepared at Washington University Genome
Sequencing Center."

ORIGIN

Query Match 55.6%; Score 55.6; DB 8; Length 717;
Best Local Similarity 77.9%; Pred. No. 9.1e-09;
Matches 67; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 3 CGGATCGGAGATCTCCGATCCCTATGGTCGACTCTCAGTACAACTGCTCTGATGCC 62
DB 248 CGGATCGATAGTCTCCTCGACTAGTTATGGTCGACTCTCAGTACAACTGCTCTGATGCC 189

QY 63 GCATAGTTAAGCCAGTATCTGCTCC 88

DB 188 GCATAGTTAAGCCAGCCGACACCC 163

RESULT 5

AW409112

LOCUS

DEFINITION sal10h5 Salivary Gland Library Homo sapiens cDNA, mRNA linear EST 31-DEC-2000
ACCESSION AW409112

VERSION AW409112.1 GI:11999687

KEYWORDS

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 348)
Bergheim, A., Ogilvie, E., Arndt, S., Napier, H., Taylor, M., Lovett, M.,
Simmons, A., Hide, W. and Ramsay, M.
A high density transcript map between markers D8S550 and D8S1759 on
8p22-23, using cDNA direct selection

JOURNAL

COMMENT

Contact: Ramsey M
Department of Human Genetics
South African Institute For Medical Research
P.O.Box 1038, Johannesburg, Gauteng, 2000, South Africa
Fax: 2711 489 9226
Email: micheler@mail.saimr.wits.ac.za.

FEATURES

source

1. .348
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="Salivary Gland"
/clone_lib="Salivary Gland Library"
/note="Vector: pAMP10"

ORIGIN

Query Match 53.6%; Score 53.6; DB 2; Length 348;
Best Local Similarity 80.5%; Pred. No. 4.1e-08;
Matches 62; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 11 GAGATCTCCGATCCCTATGGTCGACTCTCAGTACAACTGCTCTGATGCCGATAGTT 70
DB 65 GCGGTATACACACCGCATATGGTCGACTCTCAGTACAACTGCTCTGATGCCGATAGTT 124

QY 71 AAGCCAGTATCTGCTCC 87

DB 125 AAGCCAGTATACACTCC 141

RESULT 6

AL715724/c

LOCUS

DEFINITION AL715724 Danio rerio embryonic inner ear subtracted cDNA Danio
rerio cDNA clone BNOAA018ZF12 5', mRNA sequence.

ACCESSION

AL715724

VERSION

AL715724.1

KEYWORDS

EST.

SOURCE

Danio rerio (zebrafish)

ORGANISM

Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.

REFERENCE

AUTHORS

Coimbra, R., Weil, D., Brottier, P., Blanchard, S., Levi, M.,
Hardelin, J.P., Weissenbach, J. and Petit, C.
A subtracted cDNA library from the zebrafish (Danio rerio)
embryonic inner ear

JOURNAL

COMMENT

Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

source

1. .343
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone_lib="BNOAA018ZF12"
/tissue_type="inner ear"
/dev_stage="embryonic"
/clone_lib="Danio rerio embryonic inner ear subtracted
cDNA"
/note="subtracted cDNA library"

ORIGIN

Query Match 53.4%; Score 53.4; DB 1; Length 343;
Best Local Similarity 84.5%; Pred. No. 4.8e-08;
Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 17 TCCCGATCCCTATGGTCGACTCTCAGTACAACTGCTCTGATGCCGATAGTTAAGCCA 76
DB 280 TTACACCGCATATGGTCGACTCTCAGTACAACTGCTCTGATGCCGATAGTTAAGCCA 221

QY 77 GTATCTGCTCC 87

DB 220 GTATACACTCC 210

RESULT 7

AL714571/c

LOCUS

DEFINITION AL714571 Danio rerio embryonic inner ear subtracted cDNA Danio
rerio cDNA clone BNOAA007ZC02 5', mRNA sequence.

ACCESSION

AL714571

VERSION

AL714571.1

KEYWORDS

EST.

SOURCE

Danio rerio (zebrafish)

ORGANISM

Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 345)
REFERENCE
AUTHORS Coimbra,R., Weil,D., Brottier,P., Blanchard,S., Levi,M.,
Hardelin,J.P., Weissenbach,J. and Petit,C.
TITLE A subtracted cDNA library from the zebrafish (Danio rerio)
embryonic inner ear
JOURNAL Unpublished (2002)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Crémieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source
1..345
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="BN0AA0072C02"
/tissue_type="inner ear"
/dev_stages="embryonic"
/clone_lib="Danio rerio embryonic inner ear subtracted
cDNA"
/note="subtracted cDNA library"
ORIGIN
Query Match 53.4%; Score 53.4; DB 1; Length 345;
Best Local Similarity 84.5%; Pred. No. 4.8e-08;
Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
Qy 17 TCCGATCCCTATGTCGACTCTCAGTACAACTGCTCTGATCGCGCATAGTTAAGCCA 76
Db 280 TTCACCGCATATGTCGACTCTCAGTACAACTGCTCTGATCGCGCATAGTTAAGCCA 221
Qy 77 GTATCTGCTCC 87
Db 220 GTATACATCC 210
RESULT 8
CK119397/c
LOCUS CK119397 761 bp mRNA linear EST 01-JUN-2004
DEFINITION 212009.p1 AtM1 Arabidopsis thaliana cDNA clone MPMPGP2011009212
5-PRIME, mRNA sequence.
ACCESSION CK119397
VERSION CK119397.1 GI:47829713
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 761)
REFERENCE
AUTHORS Feilner,T., Immink,R.G.H., Cahill,D.J. and Kersten,B.
TITLE Generation of a cDNA expression library from Arabidopsis
inflorescence meristem
JOURNAL Unpublished (2003)
COMMENT Contact: Birgit Kersten
Plant Protein Chip Group, Department Lehrach
Max-Planck-Institute for Molecular Genetics
Innestr. 73, D-14195 Berlin, Germany
Tel: +49(0)30/84131648
Fax: +49(0)30/84131128
Email: Kersten@molgen.mpg.de
Insert Length: 761 Std Error: 0.00
Plate: 212 row: O column: 9
Seq primer: PQ865,
Location/Qualifiers
1..761
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Columbia"
/db_xref="GABI:954234"
/db_xref="taxon:3702"
FEATURES
source
1..766
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Columbia"
/db_xref="GABI:953059"
/db_xref="taxon:3702"
/clone="MPMPGP2011J04207"
/tissue_type="inflorescence meristem"
/dev_stages="about one week after bolting"
/lab_host="E. coli SCS-1/pSE111"
/clone_lib="AtM1"
/note="vector: pQE-30NAST-attB (AY386205); Site_1: SalI;
https://gabi.rzpd.de"

```

```

/clone="MPMPGP2011009212"
/tissue_type="inflorescence meristem"
/dev_stages="about one week after bolting"
/lab_host="E. coli SCS-1/pSE111"
/clone_lib="AtM1"
/note="vector: pQE-30NAST-attB (AY386205); Site_1: SalI;
Site_2: NotI; About 1 week after bolting, cDNA synthesis
using SuperscriptTM-system (Invitrogen) with an
oligo(dT)-primer containing NotI restriction site and a
SalI adapter. The main library (plate numbers begin with
1) of 38,000 clones was rearrayed into the sublibrary
(plate numbers begin with 201) containing 5,000 putative
expression clones. Average insert size is 1 kb. Note: The
rearrayed sublibrary (plate numbers begin with 201) was
sequenced. Library generation and sequencing was granted
in context of GABI-LAPP; data are also accessible at
https://gabi.rzpd.de"
ORIGIN
Query Match 53.4%; Score 53.4; DB 7; Length 761;
Best Local Similarity 84.5%; Pred. No. 5.6e-08;
Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
Qy 17 TCCGATCCCTATGTCGACTCTCAGTACAACTGCTCTGATCGCGCATAGTTAAGCCA 76
Db 674 TTCACCGCATATGTCGACTCTCAGTACAACTGCTCTGATCGCGCATAGTTAAGCCA 615
Qy 77 GTATCTGCTCC 87
Db 614 GTATACATCC 604
RESULT 9
CK120360/c
LOCUS CK120360 766 bp mRNA linear EST 01-JUN-2004
DEFINITION 207j04.p1 AtM1 Arabidopsis thaliana cDNA clone MPMPGP2011J04207
5-PRIME, mRNA sequence.
ACCESSION CK120360
VERSION CK120360.1 GI:47830676
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 766)
REFERENCE
AUTHORS Feilner,T., Immink,R.G.H., Cahill,D.J. and Kersten,B.
TITLE Generation of a cDNA expression library from Arabidopsis
inflorescence meristem
JOURNAL Unpublished (2003)
COMMENT Contact: Birgit Kersten
Plant Protein Chip Group, Department Lehrach
Max-Planck-Institute for Molecular Genetics
Innestr. 73, D-14195 Berlin, Germany
Tel: +49(0)30/84131648
Fax: +49(0)30/84131128
Email: Kersten@molgen.mpg.de
Insert Length: 766 Std Error: 0.00
Plate: 207 row: J column: 4
Seq primer: PQ865,
Location/Qualifiers
1..766
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Columbia"
/db_xref="GABI:953059"
/db_xref="taxon:3702"
/clone="MPMPGP2011J04207"
/tissue_type="inflorescence meristem"
/dev_stages="about one week after bolting"
/lab_host="E. coli SCS-1/pSE111"
/clone_lib="AtM1"
/note="vector: pQE-30NAST-attB (AY386205); Site_1: SalI;
https://gabi.rzpd.de"

```


Site_2: NotI; About 1 week after bolting, cDNA synthesis using SuperscriptTM-system (Invitrogen) with an oligo(dT)-primer containing NotI restriction site and a SalI adapter. The main library (plate numbers begin with 1) of 38,000 clones was rearrayed into the sublibrary (plate numbers begin with 201) containing 5,000 putative expression clones. Average insert size is 1 kb. Note: The rearrayed sublibrary (plate numbers begin with 201) was sequenced. Library generation and sequencing was granted in context of GABI-LAPP; data are also accessible at <https://gabi.rzpd.de>

ORIGIN

Query Match 53.4%; Score 53.4; DB 7; Length 766;
Best Local Similarity 84.5%; Pred. No. 5.6e-08;
Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 17 TCCGATCCCTGATGCTCAGTACAACTCTGCTGATGCGCATAGTTAAGCCA 76
Db 679 TTCACCGCATATGGTGCACTCTCAGTACAACTCTGCTGATGCGCATAGTTAAGCCA 620
Qy 77 GTATCTGCTCC 87
Db 619 GTATACACTCC 609

RESULT 10

CK117844/7844/c 788 bp mRNA linear EST 01-JUN-2004
LOCUS 209p08.p1 AtM1 Arabidopsis thaliana cDNA clone MPMPGP2011P08209
DEFINITION 5-PRIME, mRNA sequence.

ACCESSION CK117844

VERSION CK117844.1 GI:47828160

KEYWORDS EST.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids. 1 (bases 1 to 788)

REFERENCE Feilner, T., Immink, R.G.H., Cahill, D.J. and Kersten, B.

Generation of a cDNA expression library from Arabidopsis

inflorescence meristem

Unpublished (2003)

CONTACT: Birgit Kersten

Plant Protein Chip Group, Department Lehrach

Max-Planck-Institute for Molecular Genetics

Instr. 73, D-14195 Berlin, Germany

Tel: +49(0)30/84131648

Fax: +49(0)30/84131128

Email: Kersten@molgen.mpg.de

Insert Length: 788 Std Error: 0.00

Plate: 209 row: P column: 8

Seq primer: pQE65.

Location/Qualifiers

1..788

source

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/ecotype="Columbia"

/db_xref="GABI:953578"

/db_xref="taxon:3702"

/clone="MPMPGP2011P08209"

/tissue_type="inflorescence meristem"

/dev_stage="about one week after bolting"

/lab_host="E. coli SCS-1/pSE111"

/clone_lib="AtM1"

/note="Vector: pQE-3ONAST-attB (AY386205); Site 1: SalI;

Site 2: NotI; About 1 week after bolting, cDNA synthesis

using SuperscriptTM-system (Invitrogen) with an

oligo(dT)-primer containing NotI restriction site and a

SalI adapter. The main library (plate numbers begin with

1) of 38,000 clones was rearrayed into the sublibrary

(plate numbers begin with 201) containing 5,000 putative

expression clones. Average insert size is 1 kb. Note: The rearrayed sublibrary (plate numbers begin with 201) was sequenced. Library generation and sequencing was granted in context of GABI-LAPP; data are also accessible at <https://gabi.rzpd.de>

ORIGIN

Query Match 53.4%; Score 53.4; DB 7; Length 788;
Best Local Similarity 84.5%; Pred. No. 5.7e-08;
Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 17 TCCGATCCCTGATGCTCAGTACAACTCTGCTGATGCGCATAGTTAAGCCA 76
Db 514 TTCACCGCATATGGTGCACTCTCAGTACAACTCTGCTGATGCGCATAGTTAAGCCA 455
Qy 77 GTATCTGCTCC 87
Db 454 GTATACACTCC 444

RESULT 11

CL141237/c

LOCUS CL141237

DEFINITION ISB1-118117_T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-118117,

genomic survey sequence.

ACCESSION CL141237

VERSION CL141237.1 GI:40634872

KEYWORDS GSS.

SOURCE Xenopus tropicalis (western clawed frog)

ORGANISM Xenopus tropicalis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;

Xenopodinae; Xenopus; Silurana.

1 (bases 1 to 898)

REFERENCE Kremitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,

Mardis, E. and Wilson, R.

A physical map of the xenopus tropicalis genome

Unpublished (2003)

CONTACT: Richard K Wilson

Genome Sequencing Center

Washington University School of Medicine

Email: submissions@watson.wustl.edu

Insert Length: 75000 Std Error: 0.00

Seq primer: T7 TAATACGACTCACTATAGG

Class: BAC ends

High quality sequence start: 4

High quality sequence stop: 742.

Location/Qualifiers

1..898

source

/organism="Xenopus tropicalis"

/mol_type="genomic DNA"

/db_xref="taxon:8364"

/clone="ISB1-118117"

/clone_lib="ISB1"

/note="Vector: pBelOBAC11; ISB-1 Xenopus tropicalis BAC

Library Segment 1"

ORIGIN

Query Match 53.4%; Score 53.4; DB 9; Length 898;
Best Local Similarity 84.5%; Pred. No. 5.8e-08;
Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 17 TCCGATCCCTGATGCTCAGTACAACTCTGCTGATGCGCATAGTTAAGCCA 76
Db 195 TTCACCGCATATGGTGCACTCTCAGTACAACTCTGCTGATGCGCATAGTTAAGCCA 136
Qy 77 GTATCTGCTCC 87
Db 135 GTATACACTCC 125

RESULT 12

CL140877/c

```

LOCUS      CL140877                      899 bp    DNA        linear    GSS 05-JAN-2004
DEFINITION ISB1-118B12_T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-118B12,
            genomic survey sequence.
ACCESSION  CL140877
VERSION    CL140877.1 GI:40634512
KEYWORDS   GSS.
SOURCE     Xenopus tropicalis (western clawed frog)
ORGANISM   Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE  1 (bases 1 to 899)
            Kremitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
            Mardis, E. and Wilson, R.
TITLE      A physical map of the xenopus tropicalis genome
JOURNAL    Unpublished (2003)
COMMENT    Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 75000 Std Error: 0.00
            Seq primer: T7 TAATACGACTCACTATAGGG
            Class: BAC ends
            High quality sequence start: 4
            High quality sequence stop: 681.
FEATURES   source
            Location/Qualifiers
                1..899
                /organism="Xenopus tropicalis"
                /mol_type="genomic DNA"
                /db_xref="taxon:8364"
                /clone="ISB1-118B12"
                /clone_lib="ISB1"
                /note="Vector: pBelOBAC11; ISB-1 Xenopus tropicalis BAC
                Library Segment 1"
ORIGIN
Query Match      53.4%; Score 53.4; DB 9; Length 899;
Best Local Similarity 84.5%; Pred. No. 5.8e-08;
Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
Qy      17  TCCGATCCCTATGCTGCTCAGTACAACTGCTCTGATGCCGATAGTTAAGCCA 76
Db      195  TTCACCGCATATGTCATCTCTCAGTACAACTGCTCTGATGCCGATAGTTAAGCCA 136
Qy      77  GTATCTGCTCC 87
Db      135  GTATACACTCC 125

RESULT 13
LOCUS      CL123953/c
DEFINITION ISB1-84J15_T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-84J15,
            genomic survey sequence.
ACCESSION  CL123953
VERSION    CL123953.1 GI:40617588
KEYWORDS   GSS.
SOURCE     Xenopus tropicalis (western clawed frog)
ORGANISM   Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE  1 (bases 1 to 1009)
            Kremitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
            Mardis, E. and Wilson, R.
TITLE      A physical map of the xenopus tropicalis genome
JOURNAL    Unpublished (2003)
COMMENT    Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 75000 Std Error: 0.00
            Seq primer: T7 TAATACGACTCACTATAGGG

```

```

Class: BAC ends
High quality sequence start: 167
High quality sequence stop: 324.
FEATURES   source
            Location/Qualifiers
                1..1009
                /organism="Xenopus tropicalis"
                /mol_type="genomic DNA"
                /db_xref="taxon:8364"
                /clone="ISB1-84J15"
                /clone_lib="ISB1"
                /note="Vector: pBelOBAC11; ISB-1 Xenopus tropicalis BAC
                Library Segment 1"
ORIGIN
Query Match      53.4%; Score 53.4; DB 9; Length 1009;
Best Local Similarity 84.5%; Pred. No. 5.9e-08;
Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
Qy      17  TCCGATCCCTATGCTGCTCAGTACAACTGCTCTGATGCCGATAGTTAAGCCA 76
Db      252  TTCACCGCATATGTCATCTCTCAGTACAACTGCTCTGATGCCGATAGTTAAGCCA 193
Qy      77  GTATCTGCTCC 87
Db      192  GTATACACTCC 182

RESULT 14
LOCUS      AQ914559                      814 bp    DNA        linear    GSS 02-DEC-1999
DEFINITION nhe0049W21r CUGI Rice BAC Library (ECORI) Oryza sativa (japonica
            cultivar-group) genomic clone nhe0049W21r, genomic survey
            sequence.
ACCESSION  AQ914559
VERSION    AQ914559.1 GI:6511075
KEYWORDS   GSS.
ORGANISM   Oryza sativa (japonica cultivar-group)
            Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE  1 (bases 1 to 814)
            Wing, R.A. and Dean, R.A.
            A BAC End Sequencing Framework to Sequence the Rice Genome
            Unpublished (1998)
            Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Seq primer: GGAACAGCTATGACCATG
            Class: BAC ends
            High quality sequence start: 21
            High quality sequence stop: 361.
FEATURES   source
            Location/Qualifiers
                1..814
                /organism="Oryza sativa (japonica cultivar-group)"
                /mol_type="genomic DNA"
                /cultivar="Nipponbare"
                /db_xref="taxon:39947"
                /clone="nhe0049W21r"
                /tissue_type="Leaf"
                /lab_host="E. coli DH10B"
                /clone_lib="CUGI Rice BAC Library (ECORI)"
                /note="Vector: pBACindigo; Site_1: EcorI; Site_2: EcorI;
                Rice is the most important food crop in the world. Half of
                the world population, especially those inhabiting highly
                populated areas of the humid tropics and subtropics, rely
                on rice as their primary source of carbohydrate.
                Monocotyledonous rice is a diploid plant (2n=24) with a
                haploid genome equivalent of 431 Mbp (Arumuganathan and

```

Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from *Oryza sativa*, the Nipponbare variety using *EcoRI* as the cloning enzyme. The library contains 55,296 clones with an average insert size of 121 Kb providing approximately 15 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9%. Three high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening and can be requested from the Clemson University BAC/EST Resource Center (www.genome.clemson.edu)."

ORIGIN

Query Match 53.2%; Score 53.2; DB 8; Length 814;
 Best Local Similarity 78.0%; Pred. No. 6.7e-08;
 Matches 64; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 7 TGGGAGATCTCCGATCCCTATGCTGCACTCTCAGTACATCTGCTCTGATGCCGAT 66
 |||||
 Db 279 TGGGGGATTTTCACACCGCATATGGTGCATCTCAGTACATCTGCTCTGATGCCGAT 338
 |||||

QY 67 AGTTAAGCCAGTATCTGCTCC 88
 |||||

Db 339 AGTTAAGCCAGCCCGACACC 360
 |||||

RESULT 15

BZ051815
 LOCUS jnr57d03.b1 B.oleracea001 Brassica oleracea genomic, genomic survey
 DEFINITION 675 bp DNA linear GSS 09-OCT-2002
 sequence.
 ACCESSION BZ051815
 VERSION BZ051815.1 GI:23652690
 KEYWORDS GSS.
 SOURCE Brassica oleracea
 ORGANISM Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 675)
 Delehaunty, K., Fewell, G., Fulton, L., McCombie, W.R., Miner, T.,
 Nash, W., Rabinowicz, P.D., and Wilson, R.K.
 Whole genome shotgun reads from Brassica oleracea
 Unpublished (2002)
 Contact: Richard K. Wilson
 Genome Sequencing Center
 Washington University School of Medicine
 Email: submissions@watson.wustl.edu
 Plate: jnr57 row: d column: 03
 Seq primer: -21UPPOT forward
 Class: shotgun
 High quality sequence start: 29
 High quality sequence stop: 94.
 Location/Qualifiers

FEATURES

source
 1..675
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /db_xref="taxon:3712"
 /clone_lib="B.oleracea001"
 /note="Vector: pOTw13; Whole genome shotgun library from
 flowering buds. DNA was purified from a crude nuclear
 prep using Brassica oleracea T0100DH3 buds provided by
 Thomas Osborn at the University of Wisconsin. Genomic
 DNA was provided by Pablo Rabinowicz (CSHL) and the
 shotgun library prepared at Washington University Genome
 Sequencing Center."

ORIGIN

Query Match 53.0%; Score 53; DB 8; Length 675;
 Best Local Similarity 75.6%; Pred. No. 7.6e-08;

Matches 65; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 3 CCGATCGGGAGATCTCCCGATCCCTATGGTGGACTCTCAGTACATCTGCTCTGATGCC 62
 |||||
 Db 53 CGGNACGATAGTTCCTGGACTAGTTATGGTGGACTCTCAGTACATCTGCTCTGATGCC 112
 |||||

QY 63 GCATAGTTAAGCCAGTATCTGCTCCC 88
 |||||

Db 113 GCATAGTTAAGCCAGCCCGACACC 138
 |||||

Search completed: July 14, 2005, 23:22:58
 Job time : 952.146 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw.model

Run on: July 14, 2005, 04:39:07 ; Search time 756.618 Seconds
(without alignments)
6468.225 Million cell updates/sec

Title: US-09-482-682-44_COPY_3930_4030
Perfect score: 101
Sequence: 1 ctgagtgctagagggccgct.....tccccgcgtcttccttgac 101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hg.*

3: gb_in.*

4: gb_on.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	280	6	CQ788637
2	101	100.0	4291	12	AF302189 Synthetic
3	101	100.0	5082	6	A91754 Sequence 10
4	101	100.0	5082	6	BD085110 Vertebrat
5	101	100.0	5432	6	BD234590 Screening
6	101	100.0	5432	6	AX026821 Sequence
7	101	100.0	5650	6	AX226281 Sequence
8	101	100.0	5731	6	AX202478 Sequence
9	101	100.0	5771	12	AF060226 Eukaryoti
10	101	100.0	6050	6	AX195205 Sequence
11	101	100.0	6082	6	AR278592 Sequence
12	101	100.0	6082	6	AR367288 Sequence
13	101	100.0	6082	6	AR400324 Sequence
14	101	100.0	6082	6	AR405591 Sequence
15	101	100.0	6082	6	AR563971 Sequence
16	101	100.0	6082	6	AX141045 Sequence
17	101	100.0	6082	6	AX200905 Sequence
18	101	100.0	6082	6	AX267561 Sequence
19	101	100.0	6253	6	AR031374 Sequence

20	101	100.0	6253	6	BD009742
21	101	100.0	6338	6	BD134374 Peptide 1
22	101	100.0	6338	6	AR428934 Sequence
23	101	100.0	6365	6	AX513181 Sequence
24	101	100.0	6567	6	AX128350 Sequence
25	101	100.0	6623	6	AX128345 Sequence
26	101	100.0	6639	6	AX128351 Sequence
27	101	100.0	6649	6	AX180726 Sequence
28	101	100.0	6695	6	AX128347 Sequence
29	101	100.0	6695	6	AX128353 Sequence
30	101	100.0	6695	6	AX128354 Sequence
31	101	100.0	6746	6	AX128344 Sequence
32	101	100.0	6759	6	CQ788635 Sequence
33	101	100.0	6759	6	CQ788642 Sequence
34	101	100.0	6801	6	AX128352 Sequence
35	101	100.0	6801	6	AX128355 Sequence
36	101	100.0	6818	6	AX128346 Sequence
37	101	100.0	6828	6	AX128340 Sequence
38	101	100.0	6833	6	AX128349 Sequence
39	101	100.0	6900	6	AX128341 Sequence
40	101	100.0	6956	6	AX128348 Sequence
41	101	100.0	7038	6	AX128342 Sequence
42	101	100.0	7108	6	E36262 Human semap
43	101	100.0	7108	6	AX001326 Sequence
44	101	100.0	7231	6	BD268252 Adenoviru
45	101	100.0	7427	6	CQ768745 Sequence

ALIGNMENTS

RESULT 1
LOCUS CQ788637 280 bp DNA linear PAT 24-MAR-2004
DEFINITION Sequence 117 from Patent WO2004020643.
ACCESSION CQ788637
VERSION CQ788637.1 GI:45723394
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Curtiss,R.I. and Kong,W.
TITLE Regulated bacterial lysis for gene vaccine vector delivery and antigen release
JOURNAL Patent: WO 2004020643-A 117 11-MAR-2004;
WASHINGTON UNIVERSITY (US)
FEATURES
source Location/Qualifiers
1..280
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/notes="Description of Artificial Sequence: Multiple cloning site of pYA3650"

ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 280;
Best Local Similarity 100.0%; Pred.No. 2.6e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTCGAGTCTAGAGGGCCGTTTAAACCCGCTGATCAGCCTCGACTGTCCTTAGTTGC 60
Dy 126 CTCGAGTCTAGAGGGCCGTTTAAACCCGCTGATCAGCCTCGACTGTCCTTAGTTGC 185
Qy 61 CAGCCATCTGTGTTTCCCTCCCGTCGCTTCCCTTGAC 101
Dy 186 CAGCCATCTGTGTTTCCCTCCCGTCGCTTCCCTTGAC 226

RESULT 2

LOCUS AF302189 4291 bp DNA linear SYN 11-DEC-2001
DEFINITION Synthetic construct UOATP2 fusion protein gene, complete cds.

```
ACCESSION AF302189
VERSION AF302189.1 GI:17483748
KEYWORDS synthetic construct
SOURCE other sequences; artificial sequences.
ORGANISM
REFERENCE 1 (bases 1 to 4291)
AUTHORS Zullo,S.J., Parks,W.T., Chloupkova,M., Penton,W.A., Merrill,C.R. and Eisenstadt,J.M.
TITLE Expression of oligomycin resistance (olir) in CHO cells following transfer of the mitochondrial DNA-encoded olir ATPase 6 gene to the nuclear genome
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4291)
AUTHORS Zullo,S.J., Eisenstadt,J.M., Penton,W.A. and Merrill,C.R.
TITLE Direct Submission
JOURNAL Submitted (01-SEP-2000) Lab. Biochemical Genetics, NIMH, 9000 Rockville Pike, Bethesda, MD 20892, USA
FEATURES
    source
        1..4291
            /organism="synthetic construct"
            /mol_type="genomic DNA"
            /db_xref="taxon:32630"
            /plasmid="pUOATP2"
            483..1259
            /note="contains Homo sapiens ornithine transcarbamylase leader sequence and Crictetus griseus mtATPase6 gene"
            /codon_start=1
            /transl_table=11
            /product="UOATP2 fusion protein"
            /protein_id="AAL40188.1"
            /db_xref="GI:17483749"
            /translation="MLNLRILLNNAAFRNHNPVNRFCGQPLQNNENLFSFIFTP
            TLMGLPIILIMFPFVMTSSKRLVNNRFTFOQLIKLITKQMAIHSPKGTWSL
            MLASLIIFIGTINLGLLPHFTPTTOLSMNLGMAIPWAGAVILGFRHKWDSLAHF
            LPQCTPIPLIMLVITKISLFIQPMALAVRLTANITAGHLMLHIGGATIVLTSISL
            PTAMITFILLMLILEPAVALIQAYVFTLLVSLYLDNT"
    CDS
        100.0%; Score 101; DB 12; Length 4291;
        Best Local Similarity 100.0%; Pred. No. 2.5e-20;
        Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTCGAGTCTAGAGGCCCGGTTTAAACCCGCTGATCAGCCTCGACTGTCCTTCTAGTTGC 60
Db 1299 CTCGAGTCTAGAGGCCCGGTTTAAACCCGCTGATCAGCCTCGACTGTCCTTCTAGTTGC 1358
Qy 61 CAGCCATCTGTGTTGTCCTCCCTCCCGGTGCTTCCTTGAC 101
Db 1359 CAGCCATCTGTGTTGTCCTCCCTCCCGGTGCTTCCTTGAC 1399
FEATURES
    source
        1..5082
            /organism="unidentified"
            /mol_type="genomic DNA"
            /db_xref="taxon:32644"
    ORIGIN
        Query Match 100.0%; Score 101; DB 12; Length 4291;
        Best Local Similarity 100.0%; Pred. No. 2.5e-20;
        Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
A91754 A91754 5082 bp DNA circular PAT 22-JAN-2000
LOCUS Sequence 10 from Patent WO9824810.
DEFINITION A91754
ACCESSION A91754
VERSION A91754.1 GI:6740671
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 5082)
AUTHORS Bogaert,T.A. and Deraeymaeker,M.
TITLE VERTEBRATE HOMOLOGUES OF UNC-53 PROTEIN OF C. ELEGANS
JOURNAL Patent: WO 9824810-A 10 11-JUN-1998;
JOURNAL BOGAERT THIERRY ANDRE OLIVIER (BE); DERAEMYAEKER MARC (BE)
FEATURES
    source
        1..5082
            /organism="unidentified"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32644"
    ORIGIN
        Query Match 100.0%; Score 101; DB 6; Length 5082;
        Best Local Similarity 100.0%; Pred. No. 2.5e-20;
        Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
A91754 A91754 5082 bp DNA linear PAT 27-AUG-2002
LOCUS Vertebtrate homologues of UNC-53 protein of C elegans.
DEFINITION BD085110
ACCESSION BD085110
VERSION BD085110.1 GI:22630720
KEYWORDS JP 2001522222-A/8.
SOURCE unidentified
ORGANISM unclassified
REFERENCE 1 (bases 1 to 5082)
AUTHORS Platteeuw,C.J., Arjol,C.M.B., Deraeymaeker,M., Verhasselt,P., Pujol,N.J.R., Luc, Maertens,J.S., Luyten,W., Geerts,H., Vandekerckhove,J.S., Geysen,J. and Bogaert,T.A.O.E.
TITLE Vertebtrate homologues of UNC-53 protein of C elegans
JOURNAL Patent: JP 2001522222-A 8 13-NOV-2001;
JOURNAL JANSSEN PHARMACEUTICA NV
COMMENT OS Unidentified
PN JP 2001522222-A/8
PD 13-NOV-2001
PF 03-DEC-1997 JP 1998525231
PR 04-DEC-1996 GB 9625283.8
PI CHRIST JULES PLATTEEUW,CARLOS MANUEL BUESA ARJOL,MARC PI DERAEMYAEKER,
PI PETER VERHASSELT,NATHALIE JEANNE RAYMONDE PUJOL,LUC PI JACQUES SIMON MAERTENS,
PI WALTER LUYTEN,HUGO GEERTS,JOEL STEFAAN VANDEKERCKHOVE,JOHAN PI GEYSEN,
PI THIERRY ANDRE OLIVIER EDDY BOGAERT
PC C12N15/12,C12N5/10,C12N15/85,C07K14/435,C07K16/18,A61K38/17, A61K49/00,
PC C12Q1/02,G01N33/53
CC Strandedness: Double;
CC Topology: Circular;
CC Vertebrate homologues of UNC-53 protein of C elegans
FT source
FT 1..5082
    Location/Qualifiers
        /organism="Unidentified".
FEATURES
    source
        1..5082
            /organism="unidentified"
            /mol_type="genomic DNA"
            /db_xref="taxon:32644"
    ORIGIN
        Query Match 100.0%; Score 101; DB 6; Length 5082;
        Best Local Similarity 100.0%; Pred. No. 2.5e-20;
        Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTCGAGTCTAGAGGCCCGGTTTAAACCCGCTGATCAGCCTCGACTGTCCTTCTAGTTGC 60
Db 2669 CTCGAGTCTAGAGGCCCGGTTTAAACCCGCTGATCAGCCTCGACTGTCCTTCTAGTTGC 2728
Qy 61 CAGCCATCTGTGTTGTCCTCCCTCCCGGTGCTTCCTTGAC 101
Db 2729 CAGCCATCTGTGTTGTCCTCCCTCCCGGTGCTTCCTTGAC 2769
RESULT 4
BD085110
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
```

BD234590
LOCUS BD234590 5432 bp DNA linear PAT 17-JUL-2003
DEFINITION Screening assay of Abeta-peptide.
ACCESSION BD234590
VERSION BD234590.1 GI:33044360
KEYWORDS JP 2002531141-A/2.
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 5432)
AUTHORS Peraus,G.
TITLE Screening assay of Abeta-peptide
JOURNAL Patent: JP 2002531141-A 2 24-SEP-2002;
COMMENT AVENTIS PHARMA DEUTSCHLAND GMBH
OS Artificial Sequence
PN JP 2002531141-A/2
PD 24-SEP-2002
PF 27-NOV-1999 JP 2000586944
PR 07-DEC-1998 DE 198 56 261.6
PI GISELA PERAUS
PC C12N15/09,A01K67/033,A61K45/00,A61P25/28,C12N1/15,C12N1/19, PC
C12N5/10,C12Q1/37,C12Q1/68,C12N15/00,C12N5/00 CC Description
of Artificial Sequence: Mutagen
FH Key Location/Qualifiers
FT source 1..5432
FT Location/Qualifiers
FEATURES
source
1..5432
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 5432;
Best Local Similarity 100.0%; Pred. No. 2.5e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTCGAGTCTAGAGGCCGCGTTTAAACCGCTGATCAGCTCGACTGTGCTTCTAGTTGC 60
Db 985 CTCGAGTCTAGAGGCCGCGTTTAAACCGCTGATCAGCTCGACTGTGCTTCTAGTTGC 1044
Qy 61 CAGCCATCTGTGTTTGGCCCTCCCGCTGCTTCCCTTGAC 101
Db 1045 CAGCCATCTGTGTTTGGCCCTCCCGCTGCTTCCCTTGAC 1085
RESULT 6
AX026821
LOCUS AX026821 5432 bp DNA linear PAT 16-SEP-2000
DEFINITION Sequence 9 from Patent DE19856261.
ACCESSION AX026821
VERSION AX026821.1 GI:10187947
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Peraus,G.
JOURNAL Patent: DE 19856261-C 9 30-MAR-2000;
HOBCHST MARION ROUSSEL DE GMBH (DE)
FEATURES
source
1..5432
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Mutagen"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 5432;
Best Local Similarity 100.0%; Pred. No. 2.5e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTCGAGTCTAGAGGCCGCGTTTAAACCGCTGATCAGCTCGACTGTGCTTCTAGTTGC 60
Db 985 CTCGAGTCTAGAGGCCGCGTTTAAACCGCTGATCAGCTCGACTGTGCTTCTAGTTGC 1044
Qy 61 CAGCCATCTGTGTTTGGCCCTCCCGCTGCTTCCCTTGAC 101
Db 1045 CAGCCATCTGTGTTTGGCCCTCCCGCTGCTTCCCTTGAC 1085
RESULT 7
AX226281
LOCUS AX226281 5650 bp DNA linear PAT 10-SEP-2001
DEFINITION Sequence 2 from Patent WO0161024.
ACCESSION AX226281
VERSION AX226281.1 GI:15555545
KEYWORDS Porcine circovirus
SOURCE Porcine circovirus
ORGANISM Porcine circovirus
REFERENCE 1
AUTHORS Palmer,K.E. and Pogue,G.P.
TITLE Rolling circle replicon expression vectors
JOURNAL Patent: WO 0161024-A 2 23-AUG-2001;
Large Scale Biology Corporation (US)
FEATURES
source
1..5650
/organism="Porcine circovirus"
/mol_type="unassigned DNA"
/db_xref="taxon:46221"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 5650;
Best Local Similarity 100.0%; Pred. No. 2.5e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTCGAGTCTAGAGGCCGCGTTTAAACCGCTGATCAGCTCGACTGTGCTTCTAGTTGC 60
Db 1956 CTCGAGTCTAGAGGCCGCGTTTAAACCGCTGATCAGCTCGACTGTGCTTCTAGTTGC 2015
Qy 61 CAGCCATCTGTGTTTGGCCCTCCCGCTGCTTCCCTTGAC 101
Db 2016 CAGCCATCTGTGTTTGGCCCTCCCGCTGCTTCCCTTGAC 2056
RESULT 8
AX202478
LOCUS AX202478 5731 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 66 from Patent WO0152620.
ACCESSION AX202478
VERSION AX202478.1 GI:15392206
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Barbas,C.F., Stege,J.T., Guan,X. and Dalmia,B.
TITLE Methods and compositions to modulate expression in plants
JOURNAL Patent: WO 0152620-A 66 26-JUL-2001;
The Scripps Research Institute (US); SYNGENTA AGRICULTURAL
DISCOVERY, INC. (CA)
FEATURES
source
1..5731
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="2C7-SID"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 5731;
Best Local Similarity 100.0%; Pred. No. 2.5e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTCGAGTCTAGAGGCCGCGTTTAAACCGCTGATCAGCTCGACTGTGCTTCTAGTTGC 60

```
|||||
Db 1701 CTCGAGTCTAGAGGCGCGTTTAAACCGCGTATCAGCCTCGATGCGCTTCTAGTTGC 1760
|||||
Qy 61 CAGCCATCTGTTGTTGCCCTCCCGCGTCTTCCCTTGAC 101
|||||
Db 1761 CAGCCATCTGTTGTTGCCCTCCCGCGTCTTCCCTTGAC 1801
|||||

RESULT 9
AF060226 5771 bp DNA circular SYN 14-AUG-2000
LOCUS Eukaryotic expression vector pCR3.lmbCL-XL, complete sequence.
DEFINITION Eukaryotic expression vector pCR3.lmbCL-XL
ACCESSION AF060226
VERSION AF060226.1 GI:3108232
KEYWORDS
SOURCE Eukaryotic expression vector pCR3.lmbCL-XL
ORGANISM Eukaryotic expression vector pCR3.lmbCL-XL
other sequences: artificial sequences; vectors.
REFERENCE 1 (bases 1 to 5771)
AUTHORS Pirtskhalaishvili, G., Shurin, G.V., Gambotto, A., Esche, C., Wahl, M.,
Yurkovetsky, Z.R., Robbins, P.D. and Shurin, M.R.
TITLE Transduction of dendritic cells with Bcl-xL increases their
resistance to prostate cancer-induced apoptosis and antitumor
effect in mice
J. Immunol. 165 (4), 1956-1964 (2000)
JOURNAL
MEDLINE 20384788
PUBMED 10925278
REFERENCE 2 (bases 1 to 5771)
AUTHORS Gambotto, A., Pagliano, O., Shurin, M. and Robbins, P.D.
TITLE Direct Submission
JOURNAL Submitted (17-APR-1998) Vector Core Facility, University of
Pittsburgh, 300 Technology Drive, Pittsburgh, PA 15219, USA
FEATURES
source
1. .5771
/organism="Eukaryotic expression vector pCR3.lmbCL-XL"
/mol_type="genomic DNA"
/db_xref="taxon:75965"
1. .596
/note="CMV"
638. .657
/note="T7; priming site also"
747. 1448
/note="BAUB/c form"
/codon_start=1
/product="murine BCL-XL"
/protein_id="AAC15799.1"
/db_xref="GI:3108233"
TRANSLATION="MWSQSNRELVDLFLSYKLSQKGYSSQPSDVVEENRTEAPEATEAE
RETSAINGNPSWHLASPVAWGATGHSSSLDAREVIPMAVKQALREAGDEFELRYR
RAFSDLTSQLHITPGTAYQSFQVNNELFRDGVNNGRIVAFPSFGGALCVESVDKEMQ
VLVSRIASWATYLNHLEPIWENGWMTDFVDLYGNNAASERKQERFRNWFLLTGM
TVAGVLLGLSLFSRK"
1524. .1541
/note="pCR3.1 reverse priming site"
1827. .2500
/note="ColE1"
complement (3082. .3870)
/codon_start=1
/product="neomycin/kanamycin resistance protein"
/protein_id="AAC15800.1"
/db_xref="GI:3108234"
TRANSLATION="MIFQDGLHAGSPAAWVERLFGYDWAQQTIGCSDAAVFRLSAQQR
PVLVKFDLSGALNEQDEARLSWLAITGVPCAAVLDVVTAGRDWLLLLGEVPGDL
LSHLAPAEKVSIMADAMRHLTDPATCFDHOAKHRIERARTMEAGLVDDQDQH
QGLAPAEFLARKASMPDGEDLVVTHGDACLPNTMVENGFRSGFDICGRGLGVADRYQD
IALATRDIAELGGEWADRFVLVLYGIAAPQSORIAFYRLLDFF"
3905. .4243
/note="SV40 promoter and origin"
complement (4322. .5182)
/codon_start=1
/product="ampicillin resistance protein"
/protein_id="AAC15801.1"
/db_xref="GI:3108235"

misc_feature
1827. .2500
/note="ColE1"

rep_origin
complement (3082. .3870)
/codon_start=1
/product="neomycin/kanamycin resistance protein"
/protein_id="AAC15800.1"
/db_xref="GI:3108234"

CDS
1827. .2500
/note="ColE1"

rep_origin
complement (4322. .5182)
/codon_start=1
/product="ampicillin resistance protein"
/protein_id="AAC15801.1"
/db_xref="GI:3108235"
```

```
/translation="MSIQHFRVALIIPFFAAFLCPVFAHPETLVKVKDAEDQLGARVGY
IEDLSKGLVESFREERPPMMSTFKVLLCGAVLSRIDAGQQLGRIIHYSDNLVE
YSPVTEKHLLDGMTVRELCSAAITMSDNTAANLLLTITGGPKELTAFLLHNGDHTVRL
DRWPELNEAIPIRDERITMPVAMATTLKLLTGLLTSLASRQQLIDMDEADKVGAPL
LRSLAPAGWFIADKSGRGRSGRIIAALGPDGKPSRIIVVIYITGSGQATMDERNQIA
EIGASLIKH"
5313. .5769
/note="f1"

rep_origin
complement (100.0%; Score 101; DB 12; Length 5771;
Best Local Similarity 100.0%; Pred. No. 2.5e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTCGAGTCTAGAGGCGCGTTTAAACCGCGTATCAGCCTCGATGCGCTTCTAGTTGC 60
|||||
Db 1487 CTCGAGTCTAGAGGCGCGTTTAAACCGCGTATCAGCCTCGATGCGCTTCTAGTTGC 1546
|||||
Qy 61 CAGCCATCTGTTGTTGCCCTCCCGCGTCTTCCCTTGAC 101
|||||
Db 1547 CAGCCATCTGTTGTTGCCCTCCCGCGTCTTCCCTTGAC 1587
|||||

RESULT 10
AX195205 6050 bp DNA linear PAT 28-AUG-2001
LOCUS AX195205
DEFINITION Sequence 10 from Patent WO0151626.
ACCESSION AX195205
VERSION AX195205.1 GI:15385768
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE 1
AUTHORS Lu, X., Sun, L. and Zhang, Y.
TITLE Novel plasmid dna vectors
JOURNAL Patent: WO 0151626-A 10 19-JUL-2001;
EliM BIOPHARMACEUTICALS, INC. (US)
FEATURES
source
1. .6050
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="commercial plasmid"

ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 6050;
Best Local Similarity 100.0%; Pred. No. 2.5e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTCGAGTCTAGAGGCGCGTTTAAACCGCGTATCAGCCTCGATGCGCTTCTAGTTGC 60
|||||
Db 3837 CTCGAGTCTAGAGGCGCGTTTAAACCGCGTATCAGCCTCGATGCGCTTCTAGTTGC 3896
|||||
Qy 61 CAGCCATCTGTTGTTGCCCTCCCGCGTCTTCCCTTGAC 101
|||||
Db 3897 CAGCCATCTGTTGTTGCCCTCCCGCGTCTTCCCTTGAC 3937
|||||

RESULT 11
AR278592 6082 bp DNA linear PAT 10-APR-2003
LOCUS AR278592
DEFINITION Sequence 535 from patent US 6512094.
ACCESSION AR278592
VERSION AR278592.1 GI:29712838
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 6082)
AUTHORS Xu, J., Dillon, D. C., Mitcham, J. L., Harlocker, S. L., Jiang, Y.,
Kalos, M. D., Fanger, G. R., Retter, M. W., Stolk, J. A., Day, C. H.,
Vedvick, T. S., Carter, D., Li, S. X., Wang, A., Skeiky, Y. A. W.,
```


Hepler,W.T. and Henderson,R.A.
Compositions and methods for the therapy and diagnosis of prostate cancer
JOURNAL Patent: US 6512094-A 535 28-JAN-2003;
FEATURES Location/Qualifiers
source 1..6082
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 6082;
Best Local Similarity 100.0%; Pred. No. 2.5e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCGAGTCTAGAGGCCCGGTTTAAACCGGCTGATCAGCCTCGACTGTGCTTCTAGTTGC 60
|||||
Db 5928 CTCGAGTCTAGAGGCCCGGTTTAAACCGGCTGATCAGCCTCGACTGTGCTTCTAGTTGC 5987

QY 61 CAGCCATCTGTTGTTGGCCCTCCCGGCTGCTTCCCTTGAC 101
|||||
Db 5988 CAGCCATCTGTTGTTGGCCCTCCCGGCTGCTTCCCTTGAC 6028

RESULT 12
AR367288
LOCUS AR367288 6082 bp DNA linear PAT 12-SEP-2003
DEFINITION Sequence 535 from patent US 6329505.
ACCESSION AR367288
VERSION AR367288.1 GI:34600263
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE
1 (bases 1 to 6082)
AUTHORS Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Yuguu,J.,
Reed,S.G., Kalos,M.D., Fanger,G.R., Retter,M.W., Stolk,J.A. and
Day,C.H.

TITLE Compositions and methods for therapy and diagnosis of prostate

JOURNAL Patent: US 6329505-A 535 11-DEC-2001;
FEATURES Location/Qualifiers

source 1..6082
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 6082;
Best Local Similarity 100.0%; Pred. No. 2.5e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCGAGTCTAGAGGCCCGGTTTAAACCGGCTGATCAGCCTCGACTGTGCTTCTAGTTGC 60
|||||
Db 5928 CTCGAGTCTAGAGGCCCGGTTTAAACCGGCTGATCAGCCTCGACTGTGCTTCTAGTTGC 5987

QY 61 CAGCCATCTGTTGTTGGCCCTCCCGGCTGCTTCCCTTGAC 101
|||||
Db 5988 CAGCCATCTGTTGTTGGCCCTCCCGGCTGCTTCCCTTGAC 6028

RESULT 13
AR400324
LOCUS AR400324 6082 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 535 from patent US 6620922.
ACCESSION AR400324
VERSION AR400324.1 GI:40143590
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE
1 (bases 1 to 6082)
AUTHORS Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
Kalos,M.D., Fanger,G.R., Retter,M.W., Stolk,J.A., Day,C.H.,
Vedvick,T.S., Carter,D., Li,S.X., Wang,A., Skeiky,Y.A.W.,

Hepler,W.T. and Henderson,R.A.
Compositions and methods for the therapy and diagnosis of prostate cancer
JOURNAL Patent: US 6620922-A 535 16-SEP-2003;
FEATURES Location/Qualifiers
source 1..6082
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 6082;
Best Local Similarity 100.0%; Pred. No. 2.5e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCGAGTCTAGAGGCCCGGTTTAAACCGGCTGATCAGCCTCGACTGTGCTTCTAGTTGC 60
|||||
Db 5928 CTCGAGTCTAGAGGCCCGGTTTAAACCGGCTGATCAGCCTCGACTGTGCTTCTAGTTGC 5987

QY 61 CAGCCATCTGTTGTTGGCCCTCCCGGCTGCTTCCCTTGAC 101
|||||
Db 5988 CAGCCATCTGTTGTTGGCCCTCCCGGCTGCTTCCCTTGAC 6028

RESULT 14
AR405591
LOCUS AR405591 6082 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 535 from patent US 6630305.
ACCESSION AR405591
VERSION AR405591.1 GI:40154428
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE
1 (bases 1 to 6082)
AUTHORS Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
Kalos,M.D., Fanger,G.R., Retter,M.W., Stolk,J.A., Day,C.H.,
Vedvick,T.S., Carter,D., Li,S.X., Wang,A., Skeiky,Y.A.W.,
Hepler,W.T. and Henderson,R.A.

TITLE Compositions and methods for the therapy and diagnosis of prostate

JOURNAL Patent: US 6630305-A 535 07-OCT-2003;
FEATURES Location/Qualifiers

source 1..6082
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 6082;
Best Local Similarity 100.0%; Pred. No. 2.5e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCGAGTCTAGAGGCCCGGTTTAAACCGGCTGATCAGCCTCGACTGTGCTTCTAGTTGC 60
|||||
Db 5928 CTCGAGTCTAGAGGCCCGGTTTAAACCGGCTGATCAGCCTCGACTGTGCTTCTAGTTGC 5987

QY 61 CAGCCATCTGTTGTTGGCCCTCCCGGCTGCTTCCCTTGAC 101
|||||
Db 5988 CAGCCATCTGTTGTTGGCCCTCCCGGCTGCTTCCCTTGAC 6028

RESULT 15
AR563971
LOCUS AR563971 6082 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 535 from patent US 6759515.
ACCESSION AR563971
VERSION AR563971.1 GI:53979022
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE
1 (bases 1 to 6082)
AUTHORS Xu,J., Dillon,D.C., Mitcham,J.L., Harlocker,S.L., Jiang,Y.,
Kalos,M.D., Fanger,G.R., Retter,M.W., Stolk,J.A., Day,C.H.,

Vedvick, T.S., Carter, D., Li, S.X., Wang, A., Skeiky, Y.A.W.,
 Hepler, W.T. and Henderson, R.A.
 Cancer
 Compositions and methods for the therapy and diagnosis of prostate

JOURNAL Patent: US 6759515-A 535 06-JUL-2004;

FEATURES Location/Qualifiers

source

1..6082

/organism="unknown"

/mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 6082;
 Best Local Similarity 100.0%; Pred. No. 2.5e-20;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 CTCGAGTCTAGAGGCCCGTTAAACCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGC 60
 Db 5928 CTCGAGTCTAGAGGCCCGTTAAACCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGC 5987
 Qy 61 CAGCATCTGTGTTGCCCCCTCCCGCTGCTTCTTGAC 101
 Db 5988 CAGCATCTGTGTTGCCCCCTCCCGCTGCTTCTTGAC 6028

Search completed: July 14, 2005, 14:03:32
 Job time : 758.618 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:35:42 ; Search time 142.398 Seconds
(without alignments)
4198.742 Million cell updates/sec

Title: US-09-482-682-44_COPY_3930_4030

Perfect score: 101

Sequence: 1 ctcgagctctagagggcccg.....tccccgcgtctctcttgac 101

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_16Dec04:*

1: Geneseq1980s:*

2: Geneseq1990s:*

3: Geneseq2000s:*

4: Geneseq2001as:*

5: Geneseq2001bs:*

6: Geneseq2002as:*

7: Geneseq2002bs:*

8: Geneseq2003as:*

9: Geneseq2003bs:*

10: Geneseq2003cs:*

11: Geneseq2003ds:*

12: Geneseq2004as:*

13: Geneseq2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES.

Result No.	Score	Query Match	Length	ID	Description
1	101	100.0	278	12	Adm76942 Multiple
2	101	100.0	880	13	AdS74212 Interleuk
3	101	100.0	1340	8	Acc62322 Human NOV
4	101	100.0	1353	10	Adc26320 Human NOV
5	101	100.0	1353	12	Adm35637 Novel hum
6	101	100.0	1353	12	Ado42484 Human NOV
7	101	100.0	1420	8	Acc62251 Human NOV
8	101	100.0	1461	8	Acc62237 Human NOV
9	101	100.0	1733	12	Ado42324 Human NOV
10	101	100.0	1770	10	Adj94793 Novel NOV
11	101	100.0	1772	10	Adj94791 Novel NOV
12	101	100.0	1772	10	Adj94795 Novel NOV
13	101	100.0	1782	12	Adm41037 Cytomegal
14	101	100.0	1822	10	Adc19336 cDNA enco
15	101	100.0	1822	10	Adc19334 cDNA enco
16	101	100.0	2017	8	Ado5887 Human NOV
17	101	100.0	2017	12	Adn63050 Human NOV
18	101	100.0	2017	12	Ado42504 Human NOV
19	101	100.0	2017	12	Ado42500 Human NOV
20	101	100.0	2022	12	Ado42506 Human NOV

21	101	100.0	3482	2	ADH11353	Adh11353 Vertebrat
22	101	100.0	3972	10	ADB33520	Adb33520 Plasmid p
23	101	100.0	4291	3	AAA75084	Aaa75084 Complete
24	101	100.0	5015	10	ADB33528	Adb33528 Expressio
25	101	100.0	5082	2	ADH11417	Adh11417 Plasmid p
26	101	100.0	5432	3	AAZ89476	Aaz89476 Transgeni
27	101	100.0	5650	4	AAH74866	Aah74866 Nucleocid
28	101	100.0	5650	8	ABX94356	Abx94356 Rolling c
29	101	100.0	5731	4	AAD11615	Aad11615 Six finger
30	101	100.0	5821	12	ADM97787	Adm97787 Gal4-DBD
31	101	100.0	6050	5	AAD10237	Aad10237 Commercia
32	101	100.0	6082	4	AAH93828	Aah93828 Human pro
33	101	100.0	6082	4	AAS63921	Aas63921 Human pro
34	101	100.0	6082	4	AAH85142	Aah85142 Human pro
35	101	100.0	6082	5	ACA59729	Aca59729 Prostate
36	101	100.0	6082	6	ABL95292	Ab195292 Human P51
37	101	100.0	6082	8	AAD56212	Ad56212 Human AB-
38	101	100.0	6082	8	AAD56211	Ad56211 Human AB-
39	101	100.0	6082	8	AAD56210	Ad56210 Human AB-
40	101	100.0	6082	8	ACC95456	Acc95456 Prostate
41	101	100.0	6082	10	ADB13985	Adb13985 Human pro
42	101	100.0	6082	10	ADG26401	Adg26401 Human pro
43	101	100.0	6085	8	AAD56213	Ad56213 Human AB-
44	101	100.0	6094	8	AAD56215	Ad56215 Human PSM
45	101	100.0	6097	8	AAD56214	Ad56214 Human AB-

ALIGNMENTS

RESULT 1

ADMT76942

ID ADM76942 standard; DNA; 278 BP.

XX

AC ADM76942;

XX

DT 03-JUN-2004 (first entry)

XX

DB Multiple cloning site of pYA3650 DNA sequence.

XX

KW host-vector system; microorganism; vaccine; delivery; immunisation;

KW poultry; coccidiosis; antibacterial; plasmid; vector; gene; ds.

XX

OS Synthetic.

XX

PN WO2004020643-A2.

XX

PD 11-MAR-2004.

XX

PF 29-AUG-2003; 2003WO-US026883.

XX

PR 01-SEP-2002; 2002US-0407522P.

XX

FA (UNIV) UNIV WASHINGTON.

XX

PI Curtiss R, Kong W;

 XX |

CC comprising: (i) a eukaryotic expression cassette comprising a eukaryotic
CC promoter sequence, a site for insertion of a gene encoding a desired gene
CC product and a polyadenylation sequence; (ii) a prokaryotic activator-
CC promoter sequence; (iii) at least one origin of replication (ori); (iv) a
CC regulatable prokaryotic promoter, which is repressible by the repressor;
CC (v) at least one essential gene that is necessary for synthesis of a
CC rigid layer of a cell wall of a prokaryote; (vi) at least one
CC transcription terminator sequence; and (vii) at least one CpG sequence
CC motif that enhances immunogenicity. Also described: (1) a microorganism
CC comprising the host-vector system; (2) a vaccine comprising the
CC microorganism; (3) a method for delivery of a nucleic acid vector and/or
CC a desired gene product to a eukaryotic host; and (4) a method of
CC immunising a poultry against coccidiosis. The host-vector system has
CC antibacterial activity. The host vector system is useful as a vaccine for
CC immunising a poultry against coccidiosis. The present sequence represents
CC a nucleotide sequence which is used in the exemplification of the present
CC invention.

SQ Sequence 278 BP; 60 A; 82 C; 65 G; 71 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 12; Length 278;
Best Local Similarity 100.0%; Pred. No. 8.7e-24;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CTCGAGTCTAGAGGCGCGTTTAAACCGCTGATCAGCTCGACTGTGCTTCTAGTTGC 60
Db 126 CTCGAGTCTAGAGGCGCGTTTAAACCGCTGATCAGCTCGACTGTGCTTCTAGTTGC 185

Oy 61 CAGCCATCTGTTGTTGCTCCCTCCCGCTGCTTCTCTTGAC 101

Db 186 CAGCCATCTGTTGTTGCTCCCTCCCGCTGCTTCTCTTGAC 226

RESULT 2

AD574212
ID AD574212 standard; DNA; 880 BP.

XX AD574212;

XX 16-DEC-2004 (first entry)

XX Interleukin-2-Fc epsilon-gamma receptor transmembrane domain fusion gene.

XX Cancer; gene therapy; vaccine; human; interleukin-2; cytokine;

KW Fc epsilon-gamma receptor; receptor; IL-2tm2; gene; ds.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..880

FT /tag= a

FT /product= "IL2-tm2"

FT /partial

FT /transl_except= (pos:17..19,aa:Xaa)

FT /transl_except= (pos:23..25,aa:Xaa)

FT /transl_except= (pos:653..655,aa:Xaa)

FT /transl_except= (pos:662..664,aa:Xaa)

FT /transl_except= (pos:743..745,aa:Xaa)

FT /transl_except= (pos:764..766,aa:Xaa)

FT /transl_except= (pos:800..802,aa:Xaa)

FT /transl_except= (pos:845..847,aa:Xaa)

FT /transl_except= (pos:878..880,aa:Xaa)

FT /note= "No start or stop codon: Xaa= any amino acid"

FN WO2004080404-A2.

XX 23-SEP-2004.

XX 08-MAR-2004; 2004WO-US007012.

XX 07-MAR-2003; 2003US-0452989P.

XX (UTAH) UNIV UTAH RES FOUND.

XX

PI Samlowski W, Adams NB, McGregor J;

XX WPI; 2004-668877/65.

DR P-PSDB; ADS74209.

XX New fusion protein comprising human interleukin-2 and a transmembrane
PT domain of a protein and enhancing the activation of cytotoxic tumor-
PT infiltrating lymphocytes within tumor, useful in preparing a composition
PT for treating cancer.

PS Claim 17; SEQ ID NO 15; 64pp; English.

XX The present is that of IL-2tm2, a fusion gene comprising a human
CC interleukin-2 (IL-2) gene joined to DNA encoding the transmembrane domain
CC of Fc epsilon-gamma. IL-2tm2 is an example of a novel fusion gene of the
CC invention for use in cancer gene therapy that comprises a cytokine gene
CC and a transmembrane domain gene. It is derived from IL-2tm ADS74208 by
CC removal of a FLAG sequence and extraneous amino acids, and was expressed
CC from vector pCDNA3.1(+/-). The fusion protein is expressed as a membrane-
CC bound cytokine which may be displayed on the surface of mammalian tumor
CC cells. It is believed that by inducing expression of IL-2 on the surface
CC of tumor cells, IL-2 will activate tumor-infiltrating lymphocytes in
CC close proximity to tumor antigens. This activation is thought to
CC increase activity of antigen-specific T cells and hence to result in
CC destruction of tumor cells expressing those antigens. Murine spindle
CC cell skin cancer RD995 cells transfected with IL-2tm fusion gene or
CC pCMV2b (empty expression vector) were implanted subcutaneously into
CC C3H/HEB mice. Mice implanted with RD995 cells transfected with IL-2tm
CC fusion gene showed reduced tumor growth compared with controls.

SQ Sequence 880 BP; 238 A; 214 C; 198 G; 226 T; 0 U; 4 Other;

Query Match 100.0%; Score 101; DB 13; Length 880;
Best Local Similarity 100.0%; Pred. No. 1.2e-23;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CTCGAGTCTAGAGGCGCGTTTAAACCGCTGATCAGCTCGACTGTGCTTCTAGTTGC 60
Db 632 CTCGAGTCTAGAGGCGCGTTTAAACCGCTGATCAGCTCGACTGTGCTTCTAGTTGC 691

Oy 61 CAGCCATCTGTTGTTGCTCCCTCCCGCTGCTTCTCTTGAC 101

Db 692 CAGCCATCTGTTGTTGCTCCCTCCCGCTGCTTCTCTTGAC 732

RESULT 3

ACC62322
ID ACC62322 standard; cDNA; 1340 BP.

XX ACC62322;

XX 23-JUN-2003 (first entry)

XX Human NOV40b encoding cDNA SEQ ID NO:173.

XX Human; NOVX; antiatherosclerotic; hypotensive; cardiant; dermatological;
KW anorectic; immunosuppressive; cytostatic; antidiabetic; antiinfertility;
KW haemostatic; antiinflammatory; antiasthmatic; anti-HIV; immunomodulator;
KW neuroprotective; nootropic; antiparkinsonian; metabolic; antilipemic;
KW gene therapy; cardiomyopathy; atherosclerosis; hypertension; scleroderma;
KW congenital heart defect; aortic stenosis; valve disease; transplantation;
KW tuberosus sclerosis; obesity; congenital adrenal hyperplasia; diabetes;
KW prostate cancer; metabolic disorder; neoplasm; lymphoma; uterus cancer;
KW fertility; haemophilia; hypercoagulation; graft versus host disease;
KW idiopathic thrombocytopenic purpura; AIDS; bronchial asthma; anorexia;
KW Crohn's disease; multiple sclerosis; infectious disease; cancer;
KW cancer-associated cachexia; Alzheimer's disease; Parkinson's disease;
KW immune disorder; haematopoietic disorder; dyslipidaemia;
KW metabolic syndrome X; gene; ss.

OS Homo sapiens.

XX

PN WO20030233001-A2.
XX 20-MAR-2003.
XX 09-SEP-2002; 2002WO-US028538.
XX 07-SEP-2001; 2001US-0318120P.
XX 07-SEP-2001; 2001US-0318184P.
XX 10-SEP-2001; 2001US-0318430P.
XX 17-SEP-2001; 2001US-0322636P.
XX 17-SEP-2001; 2001US-0322781P.
XX 17-SEP-2001; 2001US-0322816P.
XX 17-SEP-2001; 2001US-0322817P.
XX 19-SEP-2001; 2001US-0323519P.
XX 20-SEP-2001; 2001US-0323631P.
XX 20-SEP-2001; 2001US-0323636P.
XX 25-SEP-2001; 2001US-0324969P.
XX 25-SEP-2001; 2001US-0325091P.
XX 26-SEP-2001; 2001US-0324990P.
XX 14-DEC-2001; 2001US-0341144P.
XX 26-FEB-2002; 2002US-0359599P.
XX 05-MAR-2002; 2002US-0361663P.
XX 03-MAY-2002; 2002US-0377908P.
XX 17-MAY-2002; 2002US-0381483P.
XX 29-MAY-2002; 2002US-0383863P.
XX 02-JUL-2002; 2002US-0393332P.
XX 17-JUL-2002; 2002US-0396412P.
XX 13-AUG-2002; 2002US-0403517P.
XX 06-SEP-2002; 2002US-00236417.
XX (CURA-) CURAGEN CORP.
XX
XX Agee ML, Alsobrook JP, Anderson DW, Berghs C, Boldog FL;
PI Burgess CE, Casman SJ, Catterton E, Chant JS, Chaudhuri A;
PI Crabtree J, Dipippo VA, Edinger SR, Eisen AJ, Ellerman K, Yi W;
PI Gangoli EA, Gerlach VN, Giot L, Gorman L, Guo X, Gusev VY, Ji W;
PI Kekuda R, Khramtsov NV, Leach MD, Lepley DM, Li L, Liu X;
PI Maliyankar UM, Miller CE, Ooi CE, Ort T, Padigaru M, Patturajan M;
PI Pena CEA, Rieger DK, Rothenberg ME, Shenoy SG, Shimkets RA;
PI Spaderna SK, Spytek KA, Taupier RJ, Twomlow N, Vernet CAM, Voss EZ;
PI Zerhusen BD, Zhong M;
XX
XX WPI; 2003-313241/30.
DR P-PSDB; ABR54253.
XX
XX Novel human proteins and nucleic acid encoding the proteins, useful for
PT diagnosis, treatment and prevention of disorders involving the human
PT protein or nucleic acid e.g. cardiac and neurological disorders.
XX
XX Claim 20; Page 236; 460pp; English.
XX
XX The present invention describes isolated human NOVX proteins, where X is
CC 1 to 42. ACC62236 to ACC62345 encode the human NOVX proteins given in
CC ABR54167 to ABR54276. NOVX sequences have antiatherosclerotic, cardiac,
CC hypotensive, dermatological, anorectic, immunosuppressive, cytostatic,
CC antidiabetic, antiinfertility, haemostatic, antiinflammatory, anti-HIV,
CC antiaesthetic, metabolic, immunomodulator, neuroprotective, nootropic,
CC antiparkinsonian and antilipemic activities, and can be used in gene
CC therapy. NOVX proteins are useful for treating or preventing a pathology
CC associated with a NOVX protein in humans and for treating a syndrome
CC associated with the human disease. NOVX nucleic acids, proteins and
CC antibodies can be used in the treatment and diagnosis of cardiomyopathy,
CC atherosclerosis, hypertension, congenital heart defects, aortic stenosis,
CC valve disease, tuberculous sclerosis, scleroderma, obesity, transplantation,
CC congenital adrenal hyperplasia, prostate cancer, diabetes, metabolic
CC disorders, neoplasm, lymphoma, uterus cancer, fertility, haemophilia,
CC hypercoagulation, idiopathic thrombocytopenic purpura, graft versus host
CC disease, AIDS, bronchial asthma, Crohn's disease, multiple sclerosis,
CC infectious disease, anorexia, cancer-associated cachexia, cancer,
CC Alzheimer's disease, Parkinson's disease, immune disorders,
CC haematopoietic disorders, dyslipidaemias, and metabolic syndrome X.
CC ACC62346 to ACC62465 represent PCR primers and probes for human NOVX
CC sequences, which are used in examples from the present invention.

CC ABR54277 represents a human trypsinogen protein given in comparison with
CC the human NOV35b protein in the exemplification of the present invention
XX
XX Sequence 1340 BP; 286 A; 350 C; 332 G; 370 T; 0 U; 2 Other;
SQ
Query Match 100.0%; Score 101; DB 8; Length 1340;
Best Local Similarity 100.0%; Pred. No. 1.3e-23;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTCGAGTCTAGAGGGCCCGTTTAAACCCGCTGATCAGCTGCTGCTTCTAGTTGC 60
|||
Db 1176 CTCGAGTCTAGAGGGCCCGTTTAAACCCGCTGATCAGCTGCTGCTTCTAGTTGC 1235
|||
Qy 61 CAGCATCTGTTGTTGCTCCCTCCCGTGCCTTCCTTGAC 101
|||
Db 1236 CAGCATCTGTTGTTGCTCCCTCCCGTGCCTTCCTTGAC 1276
|||
RESULT 4
ADC26320
ID ADC26320 standard; DNA; 1353 BP.
XX
XX AC ADC26320;
XX
XX DT 18-DEC-2003 (first entry)
XX
XX DE Human NOV34d DNA.
XX
XX KW NOV; cytostatic; metabolic disorder; immune; neurodegenerative;
KW circulatory; haemopoietic; wasting; cancer; gene therapy; vaccine;
KW transgenic; human; gene; ds.
XX
XX OS Homo sapiens.
XX
XX PN WO2003004687-A2.
XX
XX PD 16-JAN-2003.
XX
XX PF 03-JUL-2002; 2002WO-US021361.
XX
XX PR 05-JUL-2001; 2001US-0303046P.
XX PR 09-JUL-2001; 2001US-0303828P.
XX PR 09-JUL-2001; 2001US-0304016P.
XX PR 11-JUL-2001; 2001US-0304502P.
XX PR 13-JUL-2001; 2001US-0305262P.
XX PR 16-JUL-2001; 2001US-0305673P.
XX PR 17-JUL-2001; 2001US-0306085P.
XX PR 24-JUL-2001; 2001US-0307536P.
XX PR 27-JUL-2001; 2001US-0308228P.
XX PR 30-JUL-2001; 2001US-030877P.
XX PR 01-AUG-2001; 2001US-0309255P.
XX PR 17-AUG-2001; 2001US-0313328P.
XX PR 12-SEP-2001; 2001US-0318711P.
XX PR 19-SEP-2001; 2001US-0323380P.
XX PR 21-SEP-2001; 2001US-0323969P.
XX PR 04-JAN-2002; 2002US-0345022P.
XX PR 04-JAN-2002; 2002US-0345038P.
XX PR 28-FEB-2002; 2002US-0361172P.
XX PR 01-MAR-2002; 2002US-0360814P.
XX PR 01-MAR-2002; 2002US-0361133P.
XX PR 01-MAR-2002; 2002US-0361147P.
XX PR 05-MAR-2002; 2002US-0361677P.
XX PR 02-APR-2002; 2002US-0363637P.
XX PR 12-APR-2002; 2002US-0372326P.
XX PR 16-APR-2002; 2002US-0372990P.
XX PR 19-APR-2002; 2002US-0373881P.
XX PR 19-APR-2002; 2002US-0373921P.
XX PR 02-JUL-2002; 2002US-00188186.
XX (CURA-) CURAGEN CORP.
XX
XX Anderson DW, Berghs C, Boldog FL, Burgess CE, Casman SJ;
PI

PI Catterton E, Edinger S, Eisen AJ, Ellerman K, Gerlach V, Gorman L;
PI Guo X, Jeffers M, Kekuda R, Li L, Malyankar UM, Miller CE;
PI Padigaru M, Patturajan M, Pena CEA, Rastelli L, Shenoy S;
PI Shimkets RA, Spaderna SK, Spytek KA, Stone DJ, Taupier RJ;
PI Vernet CAM, Voss EZ, Zhong M;
XX WPI: 2003-221607/21.
DR P-PSDB; ADC26321.
DR
XX
PT New isolated NOVX polypeptide, useful for determining the presence of, or
PT predisposition to a disease associated with altered levels of expression
PT of the polypeptide, and for treating or preventing cancer.
XX
XX
PS Claim 20; SEQ ID NO 145; 478bp; English.
XX
XX
CC The invention relates to a novel isolated NOV polypeptide. The
CC polypeptide of the invention demonstrates cytostatic activity and may be
CC used for determining the presence of, or predisposition to a disease
CC associated with altered levels of expression of the polypeptide,
CC including metabolic disorders, immune disorders, neurodegenerative
CC disorders, circulatory diseases, haemopoietic disorders, wasting diseases
CC and cancer. The polypeptide may also be utilised during gene therapy
CC procedures, vaccine development and transgenic animal production. The
CC current sequence is that of the human NOV DNA of the invention.
XX
SQ Sequence 1353 BP; 324 A; 354 C; 356 G; 319 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 10; Length 1353;
Best Local Similarity 100.0%; Pred. No. 1.3e-23;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTCGAGTCTAGAGGCCGGTTAAACCCGCTGTATCAGCCTCGACTGTGCTTCTAGTTGC 60
Db 1187 CTCGAGTCTAGAGGCCGGTTAAACCCGCTGTATCAGCCTCGACTGTGCTTCTAGTTGC 1246
Qy 61 CAGCCATCTGTGTTGTCCTCCCTCCCGCTGCTTCTTCTTGAC 101
Db 1247 CAGCCATCTGTGTTGTCCTCCCTCCCGCTGCTTCTTCTTGAC 1287
RESULT 5
ADM35637
ID ADM35637 standard; DNA; 1353 BP.
XX
AC ADM35637;
XX
DT 03-JUN-2004 (first entry)
XX
DE Novel human NOVX gene for treating diabetes and obesity.
XX
KW ds; gene; antidiabetic; anorectic; screening; insulin resistance;
KW obesity; diabetes.
XX
OS Homo sapiens.
XX
XX WO2004013347-A2.
XX
XX 12-FEB-2004.
XX
XX 06-AUG-2003; 2003WO-US024504.
XX
FR 06-AUG-2002; 2002US-0401315P.
FR 06-AUG-2002; 2002US-0401316P.
FR 06-AUG-2002; 2002US-0401627P.
FR 07-AUG-2002; 2002US-0401788P.
FR 15-AUG-2002; 2002US-0403620P.
FR 20-AUG-2002; 2002US-0404649P.
FR 20-AUG-2002; 2002US-0404674P.
FR 20-AUG-2002; 2002US-0404676P.
FR 22-AUG-2002; 2002US-0405121P.
FR 22-AUG-2002; 2002US-0405232P.
FR 23-AUG-2002; 2002US-0405400P.
FR 23-AUG-2002; 2002US-0405684P.
FR

PR 23-AUG-2002; 2002US-0405687P.
PR 23-AUG-2002; 2002US-0405698P.
PR 26-AUG-2002; 2002US-0406353P.
PR 27-AUG-2002; 2002US-0406130P.
PR 27-AUG-2002; 2002US-0406131P.
PR 03-SEP-2002; 2002US-0407919P.
PR 09-SEP-2002; 2002US-0409366P.
PR 31-OCT-2002; 2002US-0422756P.
PR 02-DEC-2002; 2002US-00307817.
XX (CURA-) CURAGEN CORP.
XX
XX Berghs C, Ellerman K, Guo X, Li L, Ort T, Rieger DK, Vernet CAM;
PI Zhong M;
XX
XX WPI: 2004-191379/18.
DR P-PSDB; ADM35638.
XX
PT New NOVX nucleic acids and polypeptides, useful in identifying compounds
PT for treating conditions such as insulin resistance, obesity and diabetes.
XX
XX Claim 1; SEQ ID NO 71; 325pp; English.
XX
CC The invention relates to novel isolated human nucleic acid molecules and
CC their encoded proteins designated NOVX proteins. The nucleic acids and
CC encoded polypeptides are useful in screening for compounds useful for
CC treating conditions such as insulin resistance, obesity or diabetes. This
CC sequence corresponds to one of the genes of the invention.
XX
SQ Sequence 1353 BP; 324 A; 354 C; 356 G; 319 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 12; Length 1353;
Best Local Similarity 100.0%; Pred. No. 1.3e-23;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTCGAGTCTAGAGGCCGGTTAAACCCGCTGTATCAGCCTCGACTGTGCTTCTAGTTGC 60
Db 1187 CTCGAGTCTAGAGGCCGGTTAAACCCGCTGTATCAGCCTCGACTGTGCTTCTAGTTGC 1246
Qy 61 CAGCCATCTGTGTTGTCCTCCCTCCCGCTGCTTCTTCTTGAC 101
Db 1247 CAGCCATCTGTGTTGTCCTCCCTCCCGCTGCTTCTTCTTGAC 1287
RESULT 6
ADO42484
ID ADO42484 standard; cDNA; 1353 BP.
XX
AC ADO42484;
XX
DT 15-JUL-2004 (first entry)
XX
DE Human NOVX polynucleotide #167.
XX
KW Human; NOVX; gene; ss; cancer; atherosclerosis; diabetes;
KW Alzheimer's disease; Parkinson's disease; graft-versus-host disease;
KW scleroderma; hypertension; haemophilia;
KW idiopathic thrombocytopenic purpura; immunodeficiency; AIDS;
KW dyslipidemia; obesity; Crohn's disease; bronchial asthma; anorexia;
KW cancer-associated cachexia; multiple sclerosis; fertility.
XX
OS Homo sapiens.
XX
XX US2004058338-A1.
XX
XX 25-MAR-2004.
XX
XX 02-DEC-2002; 2002US-00307817.
XX
XX 03-DEC-2001; 2001US-0336881P.
XX 05-DEC-2001; 2001US-0336820P.
XX 07-DEC-2001; 2001US-0338285P.
XX 07-DEC-2001; 2001US-0338318P.
XX

PR 10-DEC-2001; 2001US-0338989P.
 PR 11-DEC-2001; 2001US-0339022P.
 PR 11-DEC-2001; 2001US-0339314P.
 PR 11-DEC-2001; 2001US-0339516P.
 PR 11-DEC-2001; 2001US-0339517P.
 PR 11-DEC-2001; 2001US-0339611P.
 PR 12-DEC-2001; 2001US-0340981P.
 PR 12-DEC-2001; 2001US-0341346P.
 PR 14-DEC-2001; 2001US-0340390P.
 PR 14-DEC-2001; 2001US-0340440P.
 PR 14-DEC-2001; 2001US-0340565P.
 PR 14-DEC-2001; 2001US-0340608P.
 PR 14-DEC-2001; 2001US-0341144P.
 PR 17-DEC-2001; 2001US-0341477P.
 PR 17-DEC-2001; 2001US-0341540P.
 PR 18-DEC-2001; 2001US-0341768P.
 PR 20-DEC-2001; 2001US-0342592P.
 PR 31-DEC-2001; 2001US-0344903P.
 PR 01-FEB-2002; 2002US-0353286P.
 PR 01-FEB-2002; 2002US-0353288P.
 PR 26-FEB-2002; 2002US-0359599P.
 PR 26-FEB-2002; 2002US-0359626P.
 PR 26-FEB-2002; 2002US-0359671P.
 PR 27-FEB-2002; 2002US-0359914P.
 PR 27-FEB-2002; 2002US-0359956P.
 PR 28-FEB-2002; 2002US-0360924P.
 PR 28-FEB-2002; 2002US-0360964P.
 PR 28-FEB-2002; 2002US-0361028P.
 PR 28-FEB-2002; 2002US-0361256P.
 PR 28-FEB-2002; 2002US-0361264P.
 PR 05-MAR-2002; 2002US-0361770P.
 PR 05-MAR-2002; 2002US-0362230P.
 PR 13-MAR-2002; 2002US-0364181P.
 PR 13-MAR-2002; 2002US-0364238P.
 PR 15-MAR-2002; 2002US-0364978P.
 PR 15-MAR-2002; 2002US-0365025P.
 PR 17-APR-2002; 2002US-0373288P.
 PR 15-MAY-2002; 2002US-0380981P.
 PR 16-MAY-2002; 2002US-0381004P.
 PR 17-MAY-2002; 2002US-0381495P.
 PR 28-MAY-2002; 2002US-0383534P.
 PR 28-MAY-2002; 2002US-0383744P.
 PR 29-MAY-2002; 2002US-0383829P.
 PR 29-MAY-2002; 2002US-0384024P.
 PR 02-JUL-2002; 2002US-0393332P.
 PR 06-AUG-2002; 2002US-0401315P.
 PR 07-AUG-2002; 2002US-0401788P.
 PR 20-AUG-2002; 2002US-0404676P.
 PR 23-AUG-2002; 2002US-0405400P.
 PR 23-AUG-2002; 2002US-0405684P.
 PR 23-AUG-2002; 2002US-0405687P.
 PR 23-AUG-2002; 2002US-0405698P.
 PR 26-AUG-2002; 2002US-0406353P.

(AGEE/) AGEE M L.
 PA (ALSO/) ALSOBROOK J P.
 PA (ANDE/) ANDERSON D W.
 PA (BERG/) BERGHS C.
 PA (BOLD/) BOLDOG F L.
 PA (BURG/) BURGESS C E.
 PA (CATT/) CATTERTON E.
 PA (DIPI/) DIPIPO V A.
 PA (EDIN/) EDINGER S R.
 PA (EISE/) EISEN A.
 PA (ELLE/) ELLERMAN K.
 PA (GANG/) GANGOLLI E A.
 PA (GERL/) GERLACH V.
 PA (GORM/) GORMAN L.
 PA (ROTH/) ROTHBERG B G.
 PA (GUOX/) GUO X S.
 PA (HERR/) HERRMANN J L.
 PA (HALV/) HALVORSEN Y.
 PA (JIWW/) JI W.

PA (KEKU/) KEKUDA R.
 PA (KIRA/) KHRAMTSOV N V.
 PA (LARO/) LAROCHELLE W J.
 PA (LEPL/) LEPLEY D M.
 PA (LILL/) LI L.
 PA (MACD/) MACDOUGALL J R.
 PA (MILL/) MILLER C E.
 PA (ORTT/) ORT T.
 PA (PADI/) PADIGARU M.
 PA (PATT/) PATTURAJAN M.
 PA (PENA/) PENNA C B A.
 PA (PEYM/) PEYMAN J A.
 PA (RIEG/) RIEGER D K.
 PA (ROTH/) ROTHENBERG M E.
 PA (SHEN/) SHENOY S G.
 PA (SMIT/) SMITHSON G.
 PA (SPAD/) SPADERNA S K.
 PA (SPYT/) SPYTEK K A.
 PA (STON/) STONE D J.
 PA (TAUP/) TAUPIER R J.
 PA (VERN/) VERNET C A M.
 PA (VOSS/) VOSS E Z.
 PA (ZHON/) ZHONG M.

Agee ML, Alsobrook JP, Anderson DW, Berghs C, Boldog FL;
 Burgess CE, Catterton E, Dipippo VA, Edinger SR, Eisen A;
 Ellerman K, Gangolli EA, Gerlach V, Gorman L, Rothberg BG, Guo XS;
 Herrmann JL, Halvorsen Y, Ji W, Kekuda R, Khrantsov NV;
 Larochelle WJ, Lepley DM, Li L, Macdougall JR, Miller CE, Ort T;
 Padigar M, Patturajan M, Pena CBA, Peyman JA, Rieger DK;
 Rothenberg ME, Shenoy SG, Smithson G, Spaderna SK, Spytek KA;
 Stone DJ, Taupier RJ, Vernet CAM, Voss EZ, Zhong M;
 WPI: 2004-268786/25.
 P-PSDB: ADO42485.

New human NOVX polypeptides and nucleic acid molecules, useful for
 diagnosing, preventing or treating NOVX-associated disorder, e.g. cancer,
 atherosclerosis, diabetes, Alzheimer's disease, Parkinson's disease or
 scleroderma.

Claim 20; SEQ ID NO 333; 610pp; English.

The invention relates to human NOVX polypeptides and the polynucleotides
 encoding them. The invention also relates to antibodies specific to the
 NOVX polypeptides. The polypeptides, polynucleotides and antibodies are
 useful for manufacturing a medicament for treating a syndrome associated
 with a human disease, such as a pathology associated with the NOVX
 polypeptide. The sequences are useful for diagnosing, treating or
 preventing a NOVX-associated disorder, e.g., cancer, atherosclerosis,
 diabetes, Alzheimer's disease, Parkinson's disease, graft-versus-host
 disease, scleroderma, hypertension, haemophilia, idiopathic
 thrombocytopenic purpura, immunodeficiencies, AIDS, dyslipidemia,
 obesity, Crohn's disease, bronchial asthma, anorexia, cancer-associated
 cachexia, multiple sclerosis or fertility. The nucleic acids may be used
 as hybridisation probes, in chromosome mapping, in tissue typing, in
 preventive medicine or in pharmacogenomics. This sequence represents a
 human NOVX polynucleotide of the invention.

Sequence 1353 BP; 324 A; 355 C; 356 G; 318 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 12; Length 1353;
 Best Local Similarity 100.0%; Pred. No. 1.3e-23;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTCGAGTCTAGAGGGCCCGTTTAAACCCGCTGATCAGCTCGACTGTCTTAGTTGC 60
 Db 1187 CTCGAGTCTAGAGGGCCCGTTTAAACCCGCTGATCAGCTCGACTGTCTTAGTTGC 1246

Qy 61 CAGCCATCTGTTGTTTGCCTCCCTCCCGTCCCTTCTTGAC 101
 Db 1247 CAGCCATCTGTTGTTTGCCTCCCTCCCGTCCCTTCTTGAC 1287

RESULT 7
ACC62251
ID ACC62251 standard; cDNA; 1420 BP.
XX
XX
AC ACC62251;
AC
XX
DT 23-JUN-2003 (first entry)
XX
DE Human NOV3h encoding cDNA SEQ ID NO:31.
XX
XX Human; NOVX; antiatherosclerotic; hypotensive; cardiac; dermatological;
KW anorectic; immunosuppressive; cytostatic; antidiabetic; antiinfertility;
KW haemostatic; antiinflammatory; antiasthmatic; anti-HIV; immunomodulator;
KW neuroprotective; nontropic; antiparkinsonian; metabolic; antilipaeamic;
KW gene therapy; cardiomyopathy; atherosclerosis; hypertension; scleroderma;
KW congenital heart defect; aortic stenosis; valve disease; transplantation;
KW tuberculous sclerosis; obesity; congenital adrenal hyperplasia; diabetes;
KW prostate cancer; metabolic disorder; neoplasm; lymphoma; uterus cancer;
KW fertility; haemophilia; hypercoagulation; graft versus host disease;
KW idiopathic thrombocytopenic purpura; AIDS; bronchial asthma; anorexia;
KW Crohn's disease; multiple sclerosis; infectious disease; cancer;
KW cancer-associated cachexia; Alzheimer's disease; Parkinson's disease;
KW immune disorder; haematopoietic disorder; dyslipidaemia;
KW metabolic syndrome X; gene; ss.
XX
OS Homo sapiens.
XX
XX WO2003023001-A2.
XX
XX 20-MAR-2003.
XX
XX
XX 09-SEP-2002; 2002WO-US028538.
XX
XX 07-SEP-2001; 2001US-0318120P.
XX
XX 07-SEP-2001; 2001US-0318184P.
XX
XX 10-SEP-2001; 2001US-0318430P.
XX
XX 17-SEP-2001; 2001US-0322636P.
XX
XX 17-SEP-2001; 2001US-0322781P.
XX
XX 17-SEP-2001; 2001US-0322816P.
XX
XX 17-SEP-2001; 2001US-0322817P.
XX
XX 19-SEP-2001; 2001US-0323519P.
XX
XX 20-SEP-2001; 2001US-0323631P.
XX
XX 20-SEP-2001; 2001US-0323636P.
XX
XX 25-SEP-2001; 2001US-0324969P.
XX
XX 25-SEP-2001; 2001US-0325091P.
XX
XX 26-SEP-2001; 2001US-0324990P.
XX
XX 14-DEC-2001; 2001US-0341144P.
XX
XX 26-FEB-2002; 2002US-0359599P.
XX
XX 05-MAR-2002; 2002US-0361663P.
XX
XX 03-MAY-2002; 2002US-0377908P.
XX
XX 17-MAY-2002; 2002US-0381483P.
XX
XX 29-MAY-2002; 2002US-0383863P.
XX
XX 02-JUL-2002; 2002US-0393332P.
XX
XX 17-JUL-2002; 2002US-0396412P.
XX
XX 13-AUG-2002; 2002US-0403517P.
XX
XX 06-SEP-2002; 2002US-00236417.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Agee ML, Alsobrook JP, Anderson DW, Berghs C, Boldog FL;
PI Burgess CE, Casman SJ, Catterton E, Chant JS, Chaudhuri A;
PI Crabtree J, Dipippo VA, Edinger SR, Eisen AJ, Ellerman K;
PI Gangolli EA, Gerlach VL, Giot L, Gorman L, Guo X, Gusev VV, Ji W;
PI Kekuda R, Khramsov NV, Leach MD, Lepley DM, Li L, Liu X;
PI Maliyankar UM, Miller CE, Ooi CE, Padigaru M, Patturajan M;
PI Pena CBA, Rieger DK, Rothenberg ME, Shenoy SG, Shinkets RA, Voss EZ;
PI Spaderna SK, Spytek KA, Taupier RJ, Twomlow N, Vernet CAM, Voss EZ;
PI Zerhusen BD, Zhong M;
XX
XX WPI, 2003-313241/30.
XX
XX P-PSDB; ABR54182.
XX

PT Novel human proteins and nucleic acid encoding the proteins, useful for
PT diagnosis, treatment and prevention of disorders involving the human
XX protein or nucleic acid e.g. cardiac and neurological disorders.
XX
PS Claim 20; Page 111-112; 460pp; English.
XX
XX The present invention describes isolated human NOVX proteins, where X is
CC 1 to 42. ACC62236 to ACC62345 encode the human NOVX proteins given in
CC ABR54167 to ABR54276. NOVX sequences have antiatherosclerotic, cardiac,
CC hypotensive, dermatological, anorectic, immunosuppressive, cytostatic,
CC antidiabetic, antiinfertility, haemostatic, antiinflammatory, anti-HIV,
CC antiasthmatic, metabolic, immunomodulator, neuroprotective, nontropic,
CC antiparkinsonian and antilipaeamic activities, and can be used in gene
CC therapy. NOVX proteins are useful for treating or preventing a pathology
CC associated with a NOVX protein in humans and for treating a syndrome
CC associated with the human disease. NOVX nucleic acids, proteins and
CC antibodies can be used in the treatment and diagnosis of cardiomyopathy,
CC atherosclerosis, hypertension, congenital heart defects, aortic stenosis,
CC valve disease, tuberculous sclerosis, scleroderma, obesity, transplantation,
CC congenital adrenal hyperplasia, prostate cancer, diabetes, metabolic
CC disorders, neoplasm, lymphoma, uterus cancer, fertility, haemophilia,
CC hypercoagulation, idiopathic thrombocytopenic purpura, graft versus host
CC disease, AIDS, bronchial asthma, Crohn's disease, multiple sclerosis,
CC infectious disease, anorexia, cancer-associated cachexia, cancer,
CC Alzheimer's disease, Parkinson's disease, immune disorders,
CC haematopoietic disorders, dyslipidaemias, and metabolic syndrome X.
CC ACC62346 to ACC62465 represent PCR primers and probes for human NOVX
CC sequences, which are used in examples from the present invention.
CC ABR54277 represents a human trypsinogen protein given in comparison with
CC the human NOV35b protein in the exemplification of the present invention
XX
XX Sequence 1420 BP; 326 A; 392 C; 338 G; 364 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 101; DB 8; Length 1420;
Best Local Similarity 100.0%; Pred. No. 1.3e-23;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTCGAGTCTAGAGGCGCGGTTTAAACCCGCTGATCAGCCTCGACTGTCCTTAGTTC 60
Db 1252 CTCGAGTCTAGAGGCGCGGTTTAAACCCGCTGATCAGCCTCGACTGTCCTTAGTTC 1311
Qy 61 CAGCCATCTGTTGTCCTCCCTCCCGCTCCCTTCCTTGAC 101
Db 1312 CAGCCATCTGTTGTCCTCCCTCCCGCTCCCTTCCTTGAC 1352
RESULT 8
ACC62237
ID ACC62237 standard; cDNA; 1461 BP.
XX
XX ACC62237;
AC
XX
DT 23-JUN-2003 (first entry)
XX
DE Human NOV1b encoding cDNA SEQ ID NO:3.
XX
XX Human; NOVX; antiatherosclerotic; hypotensive; cardiac; dermatological;
KW anorectic; immunosuppressive; cytostatic; antidiabetic; antiinfertility;
KW haemostatic; antiinflammatory; antiasthmatic; anti-HIV; immunomodulator;
KW neuroprotective; nontropic; antiparkinsonian; metabolic; antilipaeamic;
KW gene therapy; cardiomyopathy; atherosclerosis; hypertension; scleroderma;
KW congenital heart defect; aortic stenosis; valve disease; transplantation;
KW tuberculous sclerosis; obesity; congenital adrenal hyperplasia; diabetes;
KW prostate cancer; metabolic disorder; neoplasm; lymphoma; uterus cancer;
KW fertility; haemophilia; hypercoagulation; graft versus host disease;
KW idiopathic thrombocytopenic purpura; AIDS; bronchial asthma; anorexia;
KW Crohn's disease; multiple sclerosis; infectious disease; cancer;
KW cancer-associated cachexia; Alzheimer's disease; Parkinson's disease;
KW immune disorder; haematopoietic disorder; dyslipidaemia;
KW metabolic syndrome X; gene; ss.
XX
XX Homo sapiens.
XX


```

PR 28-FEB-2002; 2002US-0360964P.
PR 28-FEB-2002; 2002US-0361028P.
PR 28-FEB-2002; 2002US-0361256P.
PR 28-FEB-2002; 2002US-0361264P.
PR 05-MAR-2002; 2002US-0361770P.
PR 05-MAR-2002; 2002US-0362230P.
PR 13-MAR-2002; 2002US-0364181P.
PR 13-MAR-2002; 2002US-0364238P.
PR 15-MAR-2002; 2002US-0364978P.
PR 15-MAR-2002; 2002US-0365025P.
PR 17-APR-2002; 2002US-0373288P.
PR 15-MAY-2002; 2002US-0380981P.
PR 16-MAY-2002; 2002US-0381004P.
PR 17-MAY-2002; 2002US-0381495P.
PR 28-MAY-2002; 2002US-0383534P.
PR 28-MAY-2002; 2002US-0383744P.
PR 29-MAY-2002; 2002US-0383829P.
PR 29-MAY-2002; 2002US-0384024P.
PR 02-JUL-2002; 2002US-0393332P.
PR 06-AUG-2002; 2002US-0401315P.
PR 07-AUG-2002; 2002US-0401788P.
PR 20-AUG-2002; 2002US-0404676P.
PR 23-AUG-2002; 2002US-0405400P.
PR 23-AUG-2002; 2002US-0405684P.
PR 23-AUG-2002; 2002US-0405687P.
PR 23-AUG-2002; 2002US-0405698P.
PR 26-AUG-2002; 2002US-0406353P.
XX
PA (AGEE//) AGEE M L.
PA (ALSO//) ALSOBROOK J P.
PA (ANDE//) ANDERSON D W.
PA (BERG//) BERGHS C.
PA (BOLD//) BOLDOG F L.
PA (BURG//) BURGESS C E.
PA (CATT//) CATTERTON E.
PA (DIPI//) DIPPO V A.
PA (EDIN//) EDINGER S R.
PA (EISE//) EISEN A.
PA (ELLE//) ELLERMAN K.
PA (GANG//) GANGOLLI E A.
PA (GERL//) GERLACH V.
PA (GORM//) GORMAN L.
PA (ROTH//) ROTHBERG B G.
PA (GUOX//) GUO X S.
PA (HERR//) HERRMANN J L.
PA (HALV//) HALVORSEN Y.
PA (JIWJ//) JI W.
PA (KEKU//) KEKUDA R.
PA (KHRA//) KHRAMTSOV N V.
PA (LARO//) LAROCHELLE W J.
PA (LEPL//) LEPLEY D M.
PA (LILL//) LI L.
PA (MACD//) MACDOUGALL J R.
PA (MILL//) MILLER C E.
PA (ORTT//) ORT T.
PA (PADI//) PADIGARU M.
PA (PATT//) PATTURAJAN M.
PA (PENA//) PENNA C E A.
PA (PEYM//) PEYMAN J A.
PA (RIEG//) RIEGER D K.
PA (ROTH//) ROTHENBERG M E.
PA (SHEN//) SHENOY S G.
PA (SMIT//) SMITHSON G.
PA (SPAD//) SPADERNA S K.
PA (SPYT//) SPYTEK K A.
PA (STON//) STONE D J.
PA (TAUP//) TAUPIER R J.
PA (VERN//) VERNET C A M.
PA (VOSS//) VOSS E Z.
PA (ZHON//) ZHONG M.
XX
PI Agee ML, Alsobrook JP, Anderson DW, Berghs C, Boldog FL,
PI Burgess CE, Catterton E, Dipippo VA, Edinger SR, Eisen A;

```

```

PI Ellerman K, Gangolli EA, Gerlach V, Gorman L, Rothberg BG, Guo XS;
PI Herrmann JL, Halvorsen Y, Ji W, Kekuda R, Khrantsov NV,
PI Larochele WJ, Lepley DM, Li L, Macdougall JR, Miller CE, Ort T;
PI Padigaru M, Patturajan M, Pena CEA, Peyman JA, Rieger DK;
PI Rothenberg ME, Shenoy SG, Smithson G, Spaderna SK, Spytek KA;
PI Stone DJ, Taupier RJ, Vernet CAM, Voss EZ, Zhong M;
XX
DR WPI; 2004-268786/25.
DR P-PSDB; ADO42325.
XX
PT New human NOVX polypeptides and nucleic acid molecules, useful for
PT diagnosing, preventing or treating NOVX-associated disorder, e.g. cancer,
PT atherosclerosis, diabetes, Alzheimer's disease, Parkinson's disease or
PT scleroderma.
XX
PS Claim 20; SEQ ID NO 173; 610pp; English.
XX
CC The invention relates to human NOVX polypeptides and the polynucleotides
CC encoding them. The invention also relates to antibodies specific to the
CC NOVX polypeptides. The polypeptides, polynucleotides and antibodies are
CC useful for manufacturing a medicament for treating a syndrome associated
CC with a human disease, such as a pathology associated with the NOVX
CC polypeptide. The sequences are useful for diagnosing, treating or
CC preventing a NOVX-associated disorder, e.g., cancer, atherosclerosis,
CC diabetes, Alzheimer's disease, Parkinson's disease, graft-versus-host
CC disease, scleroderma, hypertension, haemophilia, idiopathic
CC thrombocytopenic purpura, immunodeficiencies, AIDS, dyslipidemia,
CC obesity, Crohn's disease, bronchial asthma, anorexia, cancer-associated
CC cachexia, multiple sclerosis or fertility. The nucleic acids may be used
CC as hybridisation probes, in chromosome mapping, in tissue typing, in
CC preventive medicine or in pharmacogenomics. This sequence represents a
CC human NOVX polynucleotide of the invention.
XX
SQ Sequence 1733 BP; 423 A; 431 C; 460 G; 419 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 12; Length 1733;
Best Local Similarity 100.0%; Pred. No. 1.4e-23;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTCGAGTCTAGAGGGCCGTTTAAACCCGCTGATCAGCTCGACTGTCCTTAGTTGC 60
Db 1580 CTCGAGTCTAGAGGGCCGTTTAAACCCGCTGATCAGCTCGACTGTCCTTAGTTGC 1639
Qy 61 CAGCCATCTGTTGTTGCCCTCCCGCTGCCCTTCCTTGAC 101
Db 1640 CAGCCATCTGTTGTTGCCCTCCCGCTGCCCTTCCTTGAC 1680
RESULT 10
ADJ94793
ID ADJ94793 standard; DNA; 1770 BP.
XX
AC ADJ94793;
XX
DT 06-MAY-2004 (first entry)
XX
DE Novel NOVX gene sequence #11.
XX
ds; gene; antidiabetic; anorectic; cardiant; hypotensive;
antiarteriosclerotic; anorectic; virucide; antibacterial; fungicide;
protozoacide; nootropic; neuroprotective; antiparkinsonian;
anticonvulsant; osteopathic; antiarthritic; gene therapy;
dermatological; antiasthmatic; antilipemic; infectious disease; anorexia;
metabolic disorder; diabetes; obesity; infectious disease; anorexia;
cancer; cardiovascular disease; hypertension; atherosclerosis;
neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
epilepsy; immune disorder; osteoarthritis; hematopoietic disorder;
inflammatory skin disorder; asthma; dyslipidemia; neurogenesis;
cell differentiation; cell proliferation; hematopoiesis; wound healing;
angiogenesis; chromosome mapping; tissue typing; pharmacogenomic.
OS Homo sapiens.
XX

```

PN WO2003040325-A2.
XX 15-MAY-2003.
XX 05-NOV-2002; 2002WO-US035464.
XX 05-NOV-2001; 2001US-0338626P.
PR 06-NOV-2001; 2001US-0333072P.
PR 09-NOV-2001; 2001US-0348283P.
PR 15-NOV-2001; 2001US-0335610P.
PR 16-NOV-2001; 2001US-0338543P.
PR 20-NOV-2001; 2001US-0331630P.
PR 20-NOV-2001; 2001US-0331641P.
PR 21-NOV-2001; 2001US-0332152P.
PR 27-NOV-2001; 2001US-0333461P.
PR 28-NOV-2001; 2001US-0333912P.
PR 28-NOV-2001; 2001US-0334027P.
PR 29-NOV-2001; 2001US-0334300P.
PR 30-NOV-2001; 2001US-0334421P.
PR 30-NOV-2001; 2001US-0334526P.
PR 04-DEC-2001; 2001US-0336576P.
PR 04-DEC-2001; 2001US-0336664P.
PR 07-DEC-2001; 2001US-0338314P.
PR 07-DEC-2001; 2001US-0338390P.
PR 10-DEC-2001; 2001US-0339008P.
PR 11-DEC-2001; 2001US-0339008P.
PR 11-DEC-2001; 2001US-0339286P.
PR 01-FEB-2002; 2002US-0353280P.
PR 01-FEB-2002; 2002US-0353288P.
PR 04-FEB-2002; 2002US-0354392P.
PR 04-FEB-2002; 2002US-0354393P.
PR 04-FEB-2002; 2002US-0354409P.
PR 27-FEB-2002; 2002US-0359944P.
PR 27-FEB-2002; 2002US-0360148P.
PR 05-MAR-2002; 2002US-0361790P.
PR 05-MAR-2002; 2002US-0361833P.
PR 05-MAR-2002; 2002US-0361925P.
PR 05-MAR-2002; 2002US-0362230P.
PR 05-MAR-2002; 2002US-0362625P.
PR 13-MAR-2002; 2002US-0364000P.
PR 13-MAR-2002; 2002US-0364181P.
PR 13-MAR-2002; 2002US-0364182P.
PR 13-MAR-2002; 2002US-0364197P.
PR 17-MAY-2002; 2002US-0364227P.
PR 17-MAY-2002; 2002US-0381621P.
PR 17-JUL-2002; 2002US-0383675P.
PR 17-JUL-2002; 2002US-0396703P.
PR 06-AUG-2002; 2002US-0401552P.
PR 07-AUG-2002; 2002US-0401594P.
PR 07-AUG-2002; 2002US-0401787P.
PR 15-AUG-2002; 2002US-0403619P.
PR 20-AUG-2002; 2002US-0404821P.
PR 23-AUG-2002; 2002US-0405368P.
PR 23-AUG-2002; 2002US-0405402P.
PR 23-AUG-2002; 2002US-0405496P.
PR 23-AUG-2002; 2002US-0405631P.
PR 26-AUG-2002; 2002US-0406125P.
PR 04-NOV-2002; 2002US-00287226.
XX (CURA-) CURAGEN CORP.
XX Agee ML, Alsobrook JP, Berghs C, Boldog PL, Burgess CE, Chant JS;
PI Chaudhuri A, Dipippo VA, Edinger SR, Eisen A, Ellerman K;
PI Gangolli EA, Gorman L, Gerlach VL, Ji W, Kekuda R, Khrantsov NV;
PI Li L, Malyankar UM, Macdougall JR, Mezes PS, Miller CE, Millet I;
PI Ooi CE, Ort T, Padigar M, Patturajan M, Rastelli L, Rieger DK;
PI Rothenberg ME, Shenoy SG, Spaderna SK, Spytek KA, Taupier RJ;
PI Vernet CM, Zerhusen BD, Zhong M;
XX WPI; 2003-441551/41.
DR P-PSDB; ADJ94794.
XX New isolated NOVX polypeptides and polynucleotides, useful for

PT preventing, diagnosing or treating NOVX-associated disorders, e.g. osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease, asthma, or infections.
PT Claim 20; SEQ ID NO 21; 800pp; English.
XX The invention relates to novel isolated polypeptides, mature forms of these, or a sequence that is at least 95 % identical to, or having one or more conservative amino acid substitutions in the polypeptides. The polypeptides, nucleic acid molecules and antibodies are useful in the manufacture of a medicament for treating a syndrome associated with a human disease, preferably a NOVX-associated disorder. The nucleic acid molecules, polypeptides and antibodies are useful for treating, preventing or diagnosing diseases such as metabolic disorders, diabetes, obesity, infectious diseases (viral, bacterial, fungal, helminthic, and protozoal), anorexia, cancer, cardiovascular diseases (hypertension, atherosclerosis), neurodegenerative disorders, Alzheimer's disease, Parkinson's disease, epilepsy, immune disorders (osteoarthritis), hematopoietic disorders, inflammatory skin disorders, asthma, and various dyslipidemias. The nucleic acids and polypeptides may also be used as targets for the identification of small molecules that modulate or inhibit e.g. neurogenesis, cell differentiation, cell proliferation, hematopoiesis, wound healing and angiogenesis, in gene therapy, in generation of antibodies that bind immunospecifically to NOVX substances for use in therapeutic or diagnostic methods. The nucleic acids are further used as hybridization probes, in chromosome mapping, tissue typing, preventive medicine, and pharmacogenomics. This sequence corresponds to the gene encoding one of the NOVX polypeptides of the invention.
XX Sequence 1770 BP; 486 A; 427 C; 425 G; 432 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 10; Length 1770;
Best Local Similarity 100.0%; Pred. No. 1.4e-23;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTCGAGTCTAGAGGGCCGTTAAACCGCTGTATCAGCTGCTTCCTTAGTTCG 60
Db 1605 CTCGAGTCTAGAGGGCCGTTAAACCGCTGTATCAGCTGCTTCCTTAGTTCG 1664
Qy 61 CAGCCATCTGTTGTTGCTCCCTCCCGTGCCTTCCTTGAC 101
Db 1665 CAGCCATCTGTTGTTGCTCCCTCCCGTGCCTTCCTTGAC 1705
RESULT 11
ADJ94791
ID ADJ94791 standard; DNA; 1772 BP.
XX AC ADJ94791;
XX 06-MAY-2004 (first entry)
XX Novel NOVX gene sequence #10.
ds; gene; antidiabetic; anorectic; cardiant; hypotensive;
antiarteriosclerotic; anorectic; virucide; antibacterial; fungicide;
protozoicide; nootropic; neuroprotective; antiparkinsonian;
anticongulsant; osteopathic; antiarthritic; antiinflammatory;
dermatological; antiasthmatic; antilipemic; gene therapy;
metabolic disorder; diabetes; obesity; infectious disease; anorexia;
cancer; cardiovascular disease; hypertension; atherosclerosis;
neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
epilepsy; immune disorder; osteoarthritis; hematopoietic disorder;
inflammatory skin disorder; asthma; dyslipidemia; neurogenesis;
cell differentiation; cell proliferation; hematopoiesis; wound healing;
angiogenesis; chromosome mapping; tissue typing; pharmacogenomic.
XX Homo sapiens.
OS
XX WO2003040325-A2.
XX 15-MAY-2003.

```
XX PF 05-NOV-2002; 2002WO-US035464.
XX PR
XX PR 05-NOV-2001; 2001US-0338626P.
XX PR 06-NOV-2001; 2001US-0333072P.
XX PR 09-NOV-2001; 2001US-0348283P.
XX PR 15-NOV-2001; 2001US-0335610P.
XX PR 16-NOV-2001; 2001US-0338543P.
XX PR 20-NOV-2001; 2001US-0331630P.
XX PR 21-NOV-2001; 2001US-0331641P.
XX PR 21-NOV-2001; 2001US-0332152P.
XX PR 27-NOV-2001; 2001US-0333461P.
XX PR 28-NOV-2001; 2001US-0333912P.
XX PR 28-NOV-2001; 2001US-0334027P.
XX PR 29-NOV-2001; 2001US-0334300P.
XX PR 30-NOV-2001; 2001US-0334421P.
XX PR 30-NOV-2001; 2001US-0334526P.
XX PR 04-DEC-2001; 2001US-0336576P.
XX PR 07-DEC-2001; 2001US-0336664P.
XX PR 07-DEC-2001; 2001US-0338344P.
XX PR 07-DEC-2001; 2001US-0338390P.
XX PR 10-DEC-2001; 2001US-0339006P.
XX PR 10-DEC-2001; 2001US-0339008P.
XX PR 11-DEC-2001; 2001US-0339286P.
XX PR 01-FEB-2002; 2002US-0353280P.
XX PR 01-FEB-2002; 2002US-0353288P.
XX PR 04-FEB-2002; 2002US-0354392P.
XX PR 04-FEB-2002; 2002US-0354393P.
XX PR 04-FEB-2002; 2002US-0354409P.
XX PR 27-FEB-2002; 2002US-0359944P.
XX PR 27-FEB-2002; 2002US-0360148P.
XX PR 05-MAR-2002; 2002US-0361790P.
XX PR 05-MAR-2002; 2002US-0361833P.
XX PR 05-MAR-2002; 2002US-0361925P.
XX PR 05-MAR-2002; 2002US-0362230P.
XX PR 05-MAR-2002; 2002US-0362625P.
XX PR 13-MAR-2002; 2002US-0364000P.
XX PR 13-MAR-2002; 2002US-0364181P.
XX PR 13-MAR-2002; 2002US-0364182P.
XX PR 13-MAR-2002; 2002US-0364187P.
XX PR 13-MAR-2002; 2002US-0364197P.
XX PR 17-MAY-2002; 2002US-0361621P.
XX PR 28-MAY-2002; 2002US-0363675P.
XX PR 17-JUL-2002; 2002US-0396703P.
XX PR 06-AUG-2002; 2002US-0401552P.
XX PR 07-AUG-2002; 2002US-0401594P.
XX PR 07-AUG-2002; 2002US-0401787P.
XX PR 15-AUG-2002; 2002US-0403619P.
XX PR 20-AUG-2002; 2002US-0404821P.
XX PR 23-AUG-2002; 2002US-0405368P.
XX PR 23-AUG-2002; 2002US-0405402P.
XX PR 23-AUG-2002; 2002US-0405496P.
XX PR 23-AUG-2002; 2002US-0405631P.
XX PR 26-AUG-2002; 2002US-0406125P.
XX PR 04-NOV-2002; 2002US-00287226.
XX PR
XX PR (CURA-) CURAGEN CORP.
XX PR
XX PR Agee MD, Alsbrook JP, Berghs C, Boldog FL, Burgess CE, Chant JS;
XX PR Chaudhuri A, Dipippo VA, Edinger SR, Eisen A, Ellerman K;
XX PR Gangolli EA, Gorman L, Gerlach VL, Ji W, Kekuda R, Khramtsov NV;
XX PR Li L, Malyankar UM, Macdougall JR, Mezes PS, Miller CE, Millet I;
XX PR Ooi CE, Ort T, Padigaru M, Patturajan M, Rastelli L, Rieger DK;
XX PR Rothenberg MB, Shenoy SG, Spaderna SK, Spytek KA, Taupier RJ;
XX PR Vernet CAM, Zerhusen BD, Zhong M;
XX PR
XX PR WPI; 2003-441551/41.
XX PR P-PSDB; ADJ94792.
XX PR
XX PR New isolated NOVX polypeptides and polynucleotides, useful for
XX PR preventing, diagnosing or treating NOVX-associated disorders, e.g.
XX PR osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
XX PR asthma, or infections.
```

```
XX PS Claim 20; SEQ ID NO 19; 800pp; English.
XX PR
XX PR The invention relates to novel isolated polypeptides, mature forms of
XX PR these, or a sequence that is at least 95 % identical to, or having one or
XX PR more conservative amino acid substitutions in the polypeptides. The
XX PR polypeptides, nucleic acid molecules and antibodies are useful in the
XX PR manufacture of a medicament for treating a syndrome associated with a
XX PR human disease, preferably a NOVX-associated disorder. The nucleic acid
XX PR molecules, polypeptides and antibodies are useful for treating,
XX PR preventing or diagnosing diseases such as metabolic disorders, diabetes,
XX PR obesity, infectious diseases (viral, bacterial, fungal, helminthic, and
XX PR protozoal), anorexia, cancer, cardiovascular diseases (hypertension,
XX PR atherosclerosis), neurodegenerative disorders, Alzheimer's disease,
XX PR Parkinson's disease, epilepsy, immune disorders (osteoarthritis),
XX PR hematopoietic disorders, inflammatory skin disorders, asthma, and various
XX PR dyslipidemias. The nucleic acids and polypeptides may also be used as
XX PR targets for the identification of small molecules that modulate or
XX PR inhibit e.g. neurogenesis, cell differentiation, cell proliferation,
XX PR hematopoiesis, wound healing and angiogenesis, in gene therapy, in
XX PR generation of antibodies that bind immunospecifically to NOVX substances
XX PR for use in therapeutic or diagnostic methods. The nucleic acids are
XX PR further used as hybridization probes, in chromosome mapping, tissue
XX PR typing, preventive medicine, and pharmacogenomics. This sequence
XX PR corresponds to the gene encoding one of the NOVX polypeptides of the
XX PR invention.
XX PR
XX PR Sequence 1772 BP; 486 A; 428 C; 427 G; 431 T; 0 U; 0 Other;
XX PR
XX PR Query Match 100.0%; Score 101; DB 10; Length 1772;
XX PR Best Local Similarity 100.0%; Pred. No. 1.4e-23;
XX PR Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX PR
XX PR Qy 1 CTCGAGTCTAGAGGGCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCTTAGTTGC 60
XX PR | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX PR Db 1606 CTCGAGTCTAGAGGGCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCTTAGTTGC 1665
XX PR
XX PR Qy 61 CAGCCATCTGTTGTTGGCCCTCCCGGTCGCTTCCTTGAC 101
XX PR | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX PR Db 1666 CAGCCATCTGTTGTTGGCCCTCCCGGTCGCTTCCTTGAC 1706
XX PR
XX PR RESULT 12
XX PR ADJ94795
XX PR ID ADJ94795 standard; DNA; 1772 BP.
XX PR
XX PR AC ADJ94795;
XX PR
XX PR DT 06-MAY-2004 (first entry)
XX PR
XX PR XX Novel NOVX gene sequence #12.
XX PR
XX PR ds; gene; antidiabetic; anorectic; cardiatic; hypotensive;
XX PR antiarteriosclerotic; anorectic; virucide; antibacterial; fungicide;
XX PR protozoacide; nootropic; neuroprotective; antiparkinsonian;
XX PR anticonvulsant; osteopathic; antiarthritic; antiinflammatory;
XX PR dermatological; antiasthmatic; antilipemic; gene therapy;
XX PR metabolic disorder; diabetes; obesity; infectious disease; anorexia;
XX PR cancer; cardiovascular disease; hypertension; atherosclerosis;
XX PR neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
XX PR epilepsy; immune disorder; osteoarthritis; hematopoietic disorder;
XX PR inflammatory skin disorder; asthma; dyslipidemia; neurogenesis;
XX PR cell differentiation; cell proliferation; hematopoiesis; wound healing;
XX PR angiogenesis; chromosome mapping; tissue typing; pharmacogenomic.
XX PR
XX PR Homo sapiens.
XX PR
XX PR WO2003040325-A2.
XX PR
XX PR 15-MAY-2003.
XX PR
XX PR 05-NOV-2002; 2002WO-US035464.
XX PR
```

PR 05-NOV-2001; 2001US-0338628P.
PR 06-NOV-2001; 2001US-0333072P.
PR 09-NOV-2001; 2001US-0348283P.
PR 15-NOV-2001; 2001US-0335610P.
PR 16-NOV-2001; 2001US-0338543P.
PR 20-NOV-2001; 2001US-0331630P.
PR 20-NOV-2001; 2001US-0331641P.
PR 21-NOV-2001; 2001US-0332152P.
PR 27-NOV-2001; 2001US-0333461P.
PR 28-NOV-2001; 2001US-0333912P.
PR 29-NOV-2001; 2001US-0334027P.
PR 29-NOV-2001; 2001US-0334300P.
PR 30-NOV-2001; 2001US-0334421P.
PR 30-NOV-2001; 2001US-0334526P.
PR 04-DEC-2001; 2001US-0336576P.
PR 07-DEC-2001; 2001US-0336664P.
PR 07-DEC-2001; 2001US-0338314P.
PR 07-DEC-2001; 2001US-0338390P.
PR 10-DEC-2001; 2001US-0339008P.
PR 10-DEC-2001; 2001US-0339008P.
PR 11-DEC-2001; 2001US-0339286P.
PR 01-FEB-2002; 2002US-0353280P.
PR 01-FEB-2002; 2002US-0353288P.
PR 04-FEB-2002; 2002US-0354392P.
PR 04-FEB-2002; 2002US-0354393P.
PR 04-FEB-2002; 2002US-0354409P.
PR 27-FEB-2002; 2002US-0359944P.
PR 27-FEB-2002; 2002US-0360148P.
PR 05-MAR-2002; 2002US-0361790P.
PR 05-MAR-2002; 2002US-0361833P.
PR 05-MAR-2002; 2002US-0361925P.
PR 05-MAR-2002; 2002US-0362230P.
PR 13-MAR-2002; 2002US-0362625P.
PR 13-MAR-2002; 2002US-0364000P.
PR 13-MAR-2002; 2002US-0364181P.
PR 13-MAR-2002; 2002US-0364182P.
PR 13-MAR-2002; 2002US-0364197P.
PR 17-MAY-2002; 2002US-0364227P.
PR 17-MAY-2002; 2002US-0381621P.
PR 28-MAY-2002; 2002US-0383675P.
PR 17-JUL-2002; 2002US-0396703P.
PR 06-AUG-2002; 2002US-0401552P.
PR 07-AUG-2002; 2002US-0401594P.
PR 07-AUG-2002; 2002US-0401787P.
PR 15-AUG-2002; 2002US-0403619P.
PR 20-AUG-2002; 2002US-0404821P.
PR 23-AUG-2002; 2002US-0405368P.
PR 23-AUG-2002; 2002US-0405402P.
PR 23-AUG-2002; 2002US-0405496P.
PR 23-AUG-2002; 2002US-0405631P.
PR 26-AUG-2002; 2002US-0406125P.
PR 04-NOV-2002; 2002US-00287226.
XX (CURA-) CURAGEN CORP.
XX Agge ML, Alsobrook JP, Berghs C, Boldog FL, Burgess CE, Chant JS;
PI Chaudhuri A, Dipippo VA, Edinger SR, Eisen A, Ellerman K;
PI Gangolli EA, Gorman L, Gerlach VL, Kekuda R, Khrantsov NV;
PI Li L, Malyankar UM, Macdougall JR, Mezes PS, Miller CE, Millet I;
PI Ooi CE, Ort T, Padigaru M, Patturajan M, Rastelli L, Rieger DK;
PI Rothenberg ME, Shenoy SG, Spaderna SK, Taupier RJ;
PI Vernet CAM, Zerhusen BD, Zhong M;
XX WPI, 2003-441551/41.
DR P-PSDB; ADJ94796.
XX
XX New isolated NOVX polypeptides and polynucleotides, useful for
PT preventing, diagnosing or treating NOVX-associated disorders, e.g.
PT osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
PT asthma, or infections.
XX
XX Claim 20; SEQ ID NO 23; 800pp; English.

CC The invention relates to novel isolated polypeptides, mature forms of
CC these, or a sequence that is at least 95 % identical to, or having one or
CC more conservative amino acid substitutions in the polypeptides. The
CC polypeptides, nucleic acid molecules and antibodies are useful in the
CC manufacture of a medicament for treating a syndrome associated with a
CC human disease, preferably a NOVX-associated disorder. The nucleic acid
CC molecules, polypeptides and antibodies are useful for treating,
CC preventing or diagnosing diseases such as metabolic disorders, diabetes,
CC obesity, infectious diseases (viral, bacterial, fungal, helminthic, and
CC atherosclerosis), neurodegenerative disorders, Alzheimer's disease,
CC Parkinson's disease, epilepsy, immune disorders (osteoarthritis),
CC hematopoietic disorders, inflammatory skin disorders, asthma, and various
CC dyslipidemias. The nucleic acids and polypeptides may also be used as
CC targets for the identification of small molecules that modulate or
CC inhibit e.g. neurogenesis, cell differentiation, cell proliferation, in
CC hematopoiesis, wound healing and angiogenesis, in gene therapy, in
CC generation of antibodies that bind immunospecifically to NOVX substances
CC for use in therapeutic or diagnostic methods. The nucleic acids are
CC further used as hybridization probes, in chromosome mapping, tissue
CC typing, preventive medicine, and pharmacogenomics. This sequence
CC corresponds to the gene encoding one of the NOVX polypeptides of the
CC invention.
XX
SQ Sequence 1772 BP; 486 A; 428 C; 427 G; 431 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 10; Length 1772;
Best Local Similarity 100.0%; Pred. No. 1.4e-23;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTCAGTCTAGAGGCCGCTTTAAACCCGCTGATCAGCCTCGACTGCTCTAGTTGC 60
Db 1606 CTCGAGTCTAGAGGCCGCTTTAAACCCGCTGATCAGCCTCGACTGCTCTAGTTGC 1665
Qy 61 CAGCCATCTGTTGTCCTCCCTCCCGCTGCTCTCTTGAC 101
Db 1666 CAGCCATCTGTTGTCCTCCCTCCCGCTGCTCTCTTGAC 1706
RESULT 13
ADM41037
ID ADM41037 standard; DNA; 1782 BP.
XX
XX ADM41037;
XX
XX 17-JUN-2004 (first entry)
XX
XX Cytomegalovirus nucleotide sequence SEQ ID NO:5.
XX engrafting foreign replacement cell; implanting foreign replacement cell;
XX growth; differentiation; drug development; vaccine development;
XX tissue transplantation; human disease study; cytomegalovirus; gene; ds.
XX Cytomegalovirus.
XX
XX WO2004027029-A2.
XX
XX 01-APR-2004.
XX
XX 17-SEP-2003; 2003WO-US029251.
XX
XX 19-SEP-2002; 2002US-0411790P.
XX
XX (XIME-) XIMEREX INC.
XX
XX Beschornor WE, Sosa CE, Thompson SC;
XX WPI; 2004-295402/27.
XX
XX Engrafting foreign replacement cells within a fetal non-human mammal,
XX useful in producing chimeric mammals, comprises selectively destroying
XX native cells in a tissue of a fetal non-human mammal host.

PS Disclosure; SEQ ID NO 5; 48pp; English.

XX The present invention describes a method for engrafting foreign

CC replacement cells within a foetal non-human mammal, which comprises

CC selectively destroying native cells in a tissue of a foetal non-human

CC mammal host, where the number of maternal cells of the same tissue is not

CC substantially reduced, and implanting foreign replacement cells in the

CC tissue of the foetal non-human mammal host, where the foreign replacement

CC cells replace destroyed cells of the tissue. The method is useful for

CC facilitating growth and differentiation of foreign cells within a

CC mammalian host, and for producing chimeric mammals that can be used to

CC develop new drugs and vaccine, factors, drugs and tissues for

CC transplantation, also useful to study human diseases. The present

CC sequence represents a nucleotide sequence given in the Sequence Listing

CC of the present invention but not mentioned further within the

CC specification.

XX SQ Sequence 1782 BP; 458 A; 399 C; 451 G; 474 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 12; Length 1782;

Best Local Similarity 100.0%; Pred. No. 1.4e-23;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTCGAGCTAGAGGGCCGGTTAAACCCGCTGATCAGCTCGACTGTGCCTTCVAGTTGC 60

Db 1486 CTCGAGCTAGAGGGCCGGTTAAACCCGCTGATCAGCTCGACTGTGCCTTCVAGTTGC 1545

Qy 61 CAGCATCTGTTGTTTGGCCCTCCCGCTGCTTCCTTGAC 101

Db 1546 CAGCATCTGTTGTTTGGCCCTCCCGCTGCTTCCTTGAC 1586

RESULT 14

ACD19336

ID ACD19336 standard; cDNA; 1822 BP.

XX AC ACD19336;

XX 25-AUG-2003 (first entry)

DE cDNA encoding novel human protein #16.

XX Human; NOV; gene therapy; endocrine related disease; diabetes;

KW metabolism-related disease; obesity; central nervous system disorder;

KW Alzheimer's disease; Parkinson's disease; epilepsy; multiple sclerosis;

KW schizophrenia; depression; autoimmune disorder; inflammatory disorder;

KW psoriasis; allergy; lupus erythematosus; asthma; cancer;

KW inflammatory bowel disease; rheumatoid arthritis; osteoarthritis;

KW colon cancer; lung cancer; liver cancer; breast cancer; ovarian cancer;

KW prostate cancer; brain cancer; melanoma; liver disease; liver cirrhosis;

KW lung disease; emphysema; obstructive pulmonary disease; haemophilia;

KW stroke; infection; gene; ss.

OS Homo sapiens.

XX WO2003023002-A2.

PN 20-MAR-2003.

PD 09-SEP-2002; 2002WO-US028539.

PF 07-SEP-2001; 2001US-0318120P.

PR 07-SEP-2001; 2001US-0318130P.

PR 10-SEP-2001; 2001US-0318430P.

PR 17-SEP-2001; 2001US-0322636P.

PR 17-SEP-2001; 2001US-0322781P.

PR 17-SEP-2001; 2001US-0322816P.

PR 17-SEP-2001; 2001US-0322817P.

PR 19-SEP-2001; 2001US-0323519P.

PR 20-SEP-2001; 2001US-0323631P.

PR 20-SEP-2001; 2001US-0323636P.

PR 25-SEP-2001; 2001US-0324969P.

PR 25-SEP-2001; 2001US-0325091P.

PR 26-SEP-2001; 2001US-0324990P.

PR 17-APR-2002; 2002US-0373212P.

PR 06-SEP-2002; 2002US-00236177.

XX (CURA-) CURAGEN CORP.

PA Spytch KA, Patturajan M, Gorman L, Li L, Anderson DM, Zhong M;

XX Garlach VI, Vernet CAM, Ellerman K, Berghs C, Rothenberg ME, Guo X;

PI Shimkets RA, Leach MD, Catterton E, Kekuda R, Ji W, Miller CE;

PI Rieger DK, Taupier RJ, Shenoy SG, Liu X, Padigaru M, Alsobrook JP;

PI Lingley DW, Edinger SN, Burgess CE;

XX WPI: 2003-3133242/30.

DR P-PSDB; ABO14643.

XX New cytoplasmic, nuclear membrane bound or secreted polypeptides (NOVX)

PT and polynucleotides, useful in gene therapy, e.g. for treating or

PT preventing obesity, multiple sclerosis, allergy, cancers, hemophilia,

PT stroke or infections.

PS Claim 20; Page 119-120; 586pp; English.

XX The invention describes a new isolated polypeptide (NOVX). The NOVX

CC polypeptide, nucleic acid and antibody are useful as therapeutics,

CC particularly in the manufacture of a medicament for treating a syndrome

CC associated with a human disease, which includes a pathology associated

CC with NOVX polypeptide. The DNA encoding the protein is useful in gene

CC therapy for treating the disease or condition. In particular, the NOVX

CC polypeptide or polynucleotide is useful for treating endocrine/

CC metabolism-related diseases (e.g. obesity or diabetes), central nervous

CC system disorders (e.g. Alzheimer's disease, Parkinson's disease,

CC epilepsy, multiple sclerosis, schizophrenia or depression), autoimmune

CC and inflammatory disorders (e.g. psoriasis, allergy, lupus erythematosus,

CC asthma, inflammatory bowel disease, rheumatoid arthritis or

CC osteoarthritis), cancers (e.g. colon, lung, liver, breast, ovarian,

CC prostate or brain cancers, or melanoma), liver diseases (e.g. liver

CC cirrhosis), lung diseases (emphysema or obstructive pulmonary disease),

CC haemophilia, stroke, or infections (e.g. viral, bacterial or parasitic).

CC These are also useful in developing powerful assay system for functional

CC analysis of various human disorders, as well as in diagnostic

CC applications, and for monitoring the effects of drugs during clinical

CC trials. This sequence encodes a novel human NOV protein

XX SQ Sequence 1822 BP; 415 A; 552 C; 500 G; 355 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 10; Length 1822;

Best Local Similarity 100.0%; Pred. No. 1.4e-23;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTCGAGCTAGAGGGCCGGTTAAACCCGCTGATCAGCTCGACTGTGCCTTCVAGTTGC 60

Db 1656 CTCGAGCTAGAGGGCCGGTTAAACCCGCTGATCAGCTCGACTGTGCCTTCVAGTTGC 1715

Qy 61 CAGCATCTGTTGTTTGGCCCTCCCGCTGCTTCCTTGAC 101

Db 1716 CAGCATCTGTTGTTTGGCCCTCCCGCTGCTTCCTTGAC 1756

RESULT 15

ACD19334

ID ACD19334 standard; cDNA; 1822 BP.

XX AC ACD19334;

XX 25-AUG-2003 (first entry)

DE cDNA encoding novel human protein #14.

XX Human; NOV; gene therapy; endocrine related disease; diabetes;

KW metabolism-related disease; obesity; central nervous system disorder;

KW Alzheimer's disease; Parkinson's disease; epilepsy; multiple sclerosis;

KW schizophrenia; depression; autoimmune disorder; inflammatory disorder;

KW psoriasis; allergy; lupus erythematosus; asthma; cancer;

inflammatory bowel disease; rheumatoid arthritis; osteoarthritis; colon cancer; lung cancer; liver cancer; breast cancer; ovarian cancer; prostate cancer; brain cancer; melanoma; liver disease; liver cirrhosis; lung disease; emphysema; obstructive pulmonary disease; haemophilia; stroke; infection; gene; ss.

Homo sapiens.

WO2003023002-A2.

20-MAR-2003.

09-SEP-2002; 2002WO-US028539.

07-SEP-2001: 2001US-0318120P.

07-SEP-2001; 2001US-0318130P.

10-SEP-2001; 2001US-0318430P.

17-SEP-2001; 2001US-0322636P.

17-SEP-2001; 2001US-0322781P.
17-SEP-2001; 2001US-0322816P

17-SEP-2001; 2001US-0322816P.
17-SEP-2001; 2001US-0322817P.

19-SEP-2001; 2001US-0323519P.

20-SEP-2001; 2001US-0323631P.

20-SEP-2001; 2001US-0323636P.

25-SEP-2001; 2001US-0324969P.
25-SEP-2001; 2001US-0325091P

23-SEP-2001; 2001US-0323091F;
26-SEP-2001; 2001US-0324990P;

17-APR-2002; 2002US-0373212P.

06-SEP-2002; 2002US-00236177.

(CITRA-) CITRACEN CORP

Construction **Determining**

Gerlach VL, Vernet CAM, Ellerman K, Berghs C, Rothenberg ME, Guo X, Shimkets RA, Leach MD, Catterton E, Ji W, Miller CE; Rieger DK, Taupier RJ, Shenoy SG, Liou X, padigaru M, Alsobrook JP; Lepley DM, Edinger SR, Burgess CE.

WPI: 2003-313242/30.

P-PSDB; ABO14641.

New cytoplasmic, nuclear membrane bound or secreted polypeptides (NOVX) and polynucleotides, useful in gene therapy, e.g. for treating or preventing obesity, multiple sclerosis, allergy, cancers, hemophilia, stroke or infections.

Claim 20; Page 118-119; 586pp; English.

The invention describes a new isolated polypeptide (NOVX). The NOVX polypeptide, nucleic acid and antibody are useful as therapeutics, particularly in the manufacture of a medicament for treating a syndrome associated with a human disease, which includes a pathology associated with NOVX polypeptide. The DNA encoding the protein is useful in gene therapy for treating the disease or condition. In particular, the NOVX polypeptide or polynucleotide is useful for treating endocrine/metabolism-related diseases (e.g. obesity or diabetes), central nervous system disorders (e.g. Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, schizophrenia or depression), autoimmune and inflammatory disorders (e.g. psoriasis, allergy, lupus erythematosus, asthma, inflammatory bowel disease, rheumatoid arthritis or osteoarthritis), cancers (e.g. colon, lung, liver, breast, ovarian, prostate or brain cancers, or melanoma), liver diseases (e.g. liver cirrhosis), lung diseases (emphysema or obstructive pulmonary disease), haemophilia, stroke, or infections (e.g. viral, bacterial or parasitic). These are also useful in developing powerful assay system for functional analysis of various human disorders, as well as in diagnostic applications, and for monitoring the effects of drugs during clinical trials. This sequence encodes a novel human NOV protein.

Sequence 1822 BP: 415 A: 552 C: 500 G: 355 T: 0 U: 0 Other:

Query Match 100.0%; Score 101; DB 10; Length 1822;
Best Local Similarity 100.0%; Pred. No. 1.4e-23;

	Matches	101;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	CTCAGCTCTAGAGGGCCCGTTTAAACCCGCTGATCAGCTCGACTGCTGCTTCTACTTGC	60							
Db	1656	CTCAGCTCTAGAGGGCCCGTTTAAACCCGCTGATCAGCTCGACTGCTGCTTCTACTTGC	1715							
Qy	61	CAGCCATCTGTTGTTTGGCCCTCCCGCTGCTTCCCTTGAC	101							
Db	1716	CAGCCATCTGTTGTTTGGCCCTCCCGCTGCTTCCCTTGAC	1756							

Search completed: July 14, 2005, 07:01:42
Job time : 146.448 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 05:15:57 ; Search time 961.667 Seconds
(without alignments)
3997.736 Million cell updates/sec

Title: US-09-482-682-44_COPY_3930_4030

Perfect score: 101

Sequence: 1 ctccagctctagagggcccgct.....tcccocgctgccttccttgac 101

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*

1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_ges1:*
9: gb_ges2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	378	7	CF315931 HD--05-A1
2	94.6	93.7	605	7	CK719567 19817 Swo
3	77.4	76.6	400	7	CK850860 10869 Sto
4	75	74.3	295	7	CN778129 pgnzc.pk0
5	75	74.3	519	5	BM888450 TM108 Hum
6	75	74.3	521	5	BM887817 TM553 Hum
7	75	74.3	233	9	CR154962 Forward s
8	73.4	72.7	534	5	BM887701 TM304 Hum
9	73.2	72.5	132	9	CR074510 Forward s
10	73.2	72.5	329	9	CG632479 OST350781
11	72.2	71.5	286	9	CR083191 Forward s
12	70.8	70.1	130	9	BX988352 Forward s
13	69.8	69.1	75	9	CR037248 Forward s
14	69.4	68.7	600	5	BM887768 TM397 Hum
15	69.2	68.5	166	9	CR092687 Forward s
16	68.4	67.7	304	9	CR997931 Forward s
17	67	66.3	471	4	BM819796 K-EST0087
18	66.8	66.1	87	9	CR106833 Forward s
19	66.4	65.7	104	9	CR104210 Forward s
20	66.4	65.7	110	9	BX982981 Forward s
21	66.2	65.5	284	9	BX984480 Forward s
22	66	65.3	158	9	CR117924 Forward s
23	65.8	65.1	108	9	CR173214 Forward s
24	65.6	65.0	158	9	CR018574 Forward s

c 25	65.6	65.0	197	9	CR014355	CR014355 Forward s
c 26	65	64.4	323	9	CR100521	CR100521 Forward s
c 27	64	63.4	141	9	CR126132	CR126132 Forward s
c 28	64	63.4	159	9	CR133954	CR133954 Forward s
c 29	64	63.4	160	9	CR012517	CR012517 Forward s
c 30	63.8	63.2	109	9	CR108493	CR108493 Forward s
c 31	63.4	62.8	169	9	CR137375	CR137375 Forward s
c 32	63.4	62.8	234	9	CR070494	CR070494 Forward s
c 33	63.2	62.6	89	9	CR081749	CR081749 Forward s
c 34	63.2	62.6	330	9	CR006502	CR006502 Forward s
c 35	63	62.4	106	9	CR160976	CR160976 Forward s
c 36	63	62.4	605	5	BM888562	BM888562 TMM237 Hu
c 37	62.8	62.2	77	9	CR171087	CR171087 Forward s
c 38	62.8	62.2	132	9	CR081810	CR081810 Forward s
c 39	62.8	62.2	347	9	CR045655	CR045655 Forward s
c 40	62.4	61.8	141	9	CR140069	CR140069 Forward s
c 41	62.2	61.6	107	9	CR031231	CR031231 Forward s
c 42	62.2	61.6	109	9	CR065972	CR065972 Forward s
c 43	61.6	61.0	107	9	CR093214	CR093214 Forward s
c 44	61.4	60.8	113	9	CR100912	CR100912 Forward s
c 45	61.2	60.6	115	9	CR073933	CR073933 Forward s

ALIGNMENTS

RESULT 1
CF315931
LOCUS
DEFINITION HD--05-A13.b1 OSHDAC1-overexpressing transgenic rice plasmid cDNA library (HD) Oryza sativa (japonica cultivar-group) cDNA clone
ACCESSION CF315931
VERSION CF315931.1 GI:33687692
KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 378)
AUTHORS Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C., Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
TITLE Large-scale Sequencing Analysis of Rice ESTs
JOURNAL Unpublished (2003)
COMMENT Contact: Nahm B.H.
Genomics and Genetics Institute, GreenGene Biotech Inc.; Division of Bioscience and Bioinformatics, Myongji University
Yongin, Kyeonggi, Korea
Tel: 82 31 330 6193
Fax: 82 31 321 6355
Email: bnhnm@gbio.com, bnhnm@bio.myongji.ac.kr.

FEATURES
source

1. 378
Location/Qualifiers
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nackdong"
/db_xref="taxon:39947"
/clone="HD--05-A13"
/tissue_type="callus"
/dev_stage="proliferated callus on 2N6 media for 2 weeks"
/lab_host="E.coli DH108"
/clone_lib="OSHDAC1-overexpressing transgenic rice plasmid cDNA library (HD)"
/note="vector: pCR4-TOPO; Site 1: EcoRI; Callus was treated with ABA(20um) for 1hr. Oligo-capped mRNA was reverse transcribed and then used for PCR. mRNA was derived from rice Histone Deacetylase overexpression line."

ORIGIN

Query Match 100.0%; Score 101; DB 7; Length 378;
Best Local Similarity 100.0%; Pred. No. 5e-21;

```

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTCGAGTCTAGAGGCGCGGTTTAAACCGCTGATCAGCTCGACTGTCCTTAGTTGC 60
    |||||
Db 86 CTCGAGTCTAGAGGCGCGGTTTAAACCGCTGATCAGCTCGACTGTCCTTAGTTGC 145
    |||||

Qy 61 CAGCCATCTGTTGTTGCCCTCCCGCTGCTTCCTTGAC 101
    |||||
Db 146 CAGCCATCTGTTGTTGCCCTCCCGCTGCTTCCTTGAC 186
    |||||

RESULT 2
CK719567
LOCUS 19817 Swollen Stolon Solanum tuberosum cDNA, mRNA sequence.
DEFINITION CK719567
ACCESSION CK719567
VERSION CK719567.1 GI:42511281
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 605)
AUTHORS Flinn,B., Rothwell,C., Sardana,R., Griffiths,R., Lague,M., and
DeKoeyer,D., Audy,P., Goyer,C., Li,X.-Q., Wang-Pruski,G. and
Regan,S.
TITLE Generation of ESTs from swollen stolon tissues of potato
JOURNAL Unpublished (2004)
COMMENT Contact: Barry Flinn
The Canadian Potato Genome Project - BioAtlantech
921 College Hill Rd, Fredericton, ON, E3B 6Z9, CANADA
Email: bflinn@bioatlantech.nb.ca
Clones can be requested from BioAtlantech via
bflinn@bioatlantech.nb.ca
Seq primer: T3.
Location/Qualifiers
source
1..605
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Shepody"
/db_xref="taxon:4113"
/tissue_type="Stolon"
/lab_host="XL10-Gold"
/clone_lib="Swollen Stolon"
/note="Vector: pBluescript II SK(+); Site 1: EcORI;
Site 2: XhoI; supplier: Developmental series. Plants from
pathogen-free Solanum tuberosum var. Shepody, clone 1756,
nuclear stock were grown in a screenhouse under natural
conditions. RNA was isolated from swollen stolon tissue,
3-10mm in length, which was cut from the tip, to the base
of swelling."

ORIGIN
Query Match 93.7%; Score 94.6; DB 7; Length 605;
Best Local Similarity 96.0%; Pred. No. 5.5e-19;
Matches 97; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CTCGAGTCTAGAGGCGCGGTTTAAACCGCTGATCAGCTCGACTGTCCTTAGTTGC 60
    |||||
Db 400 CTCGAGTCTAGAGGCGCGGTTTAAACCGCTGATCAGCTCGACTGTCCTTAGTTGC 459
    |||||

Qy 61 CAGCCATCTGTTGTTGCCCTCCCGCTGCTTCCTTGAC 101
    |||||
Db 460 CAGCCATCTGTTGTTGCCCTCCCGCTGCTTCCTTGAC 500
    |||||

RESULT 3
CK850860
LOCUS 10869 Stolon Solanum tuberosum cDNA, mRNA sequence.
DEFINITION CK850860
ACCESSION CK850860
VERSION CK850860.1 GI:45239470

```

```

KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 400)
AUTHORS Flinn,B., Rothwell,C., Sardana,R., Griffiths,R., Lague,M., De
Koeyer,D., Audy,P., Goyer,C., Li,X.-Q., Wang-Pruski,G. and Regan,S.
TITLE Generation of ESTs from stolon tissues of potato
JOURNAL Unpublished (2004)
COMMENT Contact: Barry Flinn
The Canadian Potato Genome Project - BioAtlantech
921 College Hill Rd, Fredericton, ON, E3B 6Z9, CANADA
Email: bflinn@bioatlantech.nb.ca
Seq primer: T3.
Location/Qualifiers
source
1..400
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Shepody"
/db_xref="taxon:4113"
/tissue_type="Stolon"
/lab_host="XL10-Gold"
/clone_lib="Stolon"
/note="Vector: pBluescript II SK(+); Site 1: EcORI;
Site 2: XhoI; supplier: Developmental series. Plants from
pathogen-free Solanum tuberosum var. Shepody, clone 1756,
nuclear stock were grown in a screenhouse under natural
conditions. RNA was isolated from stolon tissue."

ORIGIN
Query Match 76.6%; Score 77.4; DB 7; Length 400;
Best Local Similarity 92.9%; Pred. No. 1.3e-13;
Matches 92; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

Qy 1 CTCGAGTCTAGAGGCGCGGTTTAAACCGCTGATCAGCTCGACTGTCCTTAGTTGC 60
    |||||
Db 231 CTCGAGTCTAGAGGCGCGGTTTAAACCGCTGATCAGCTCGACTGTCCTTAGTTGC 290
    |||||

Qy 61 CAGCCA-TCGTGTTGTTGCCCTCCCGCTGCTTCCTT 98
    |||||
Db 291 CAGCCACTCTGTTGTTGCCCTCCCGCTGCTTCCT 329
    |||||

RESULT 4
CK778129
LOCUS 295 bp mRNA linear EST 20-MAY-2004
DEFINITION pgn2c.pk001.h10.f Chicken Lymphoid cDNA library (pgn2c) Gallus
gallus cDNA clone pgn2c.pk001.h10.f 3' end of pat.pk0008.d12 5',
mRNA sequence.
ACCESSION CN778129
VERSION CN778129.1 GI:47548763
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 295)
AUTHORS Morgan,R.W. and Burnside,J.
TITLE Chicken ESTs from lymphoid tissue- 3' sequence
JOURNAL Unpublished (2004)
COMMENT Contact: Robin W. Morgan
University of Delaware
Townsend Hall, Newark, DE 19717, USA
Tel: 302-831-1341
Fax: 302-831-2822
Email: morgan@udel.edu, www.chickest.udel.edu.
Location/Qualifiers
source
1..295
/organism="Gallus gallus"
/mol_type="mRNA"

```

/db_xref="taxon:9031"
 /clone="pgn2c.pk001.h10.f 3'end of pat.pk0008.d12"
 /sex="Male and Female"
 /tissue type="thymus, bursa, spleen, PBL, bone marrow"
 /lab host="E.coli EMDH108"
 /clone_lib="Chicken lymphoid cDNA library (pgn2c)"
 /note="Vector: pcwvSPORT 6"

ORIGIN

Query Match 74.3%; Score 75; DB 7; Length 295;
 Best Local Similarity 94.0%; Pred. No. 6.8e-13;
 Matches 78; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 19 GTTTAAACCCGCTGATCAGCTCGACTGTGCTTCTAGTTGCCAGCCATCTGTTGTTGC 78
 |||
 Db 1 GCTAGAGCTCGCTGATCAGCTCGACTGTGCTTCTAGTTGCCAGCCATCTGTTGTTGC 60

Qy 79 CCTCTCCCGCTGCTTCTCTTGAC 101
 |||
 Db 61 CCTCTCCCGCTGCTTCTCTTGAC 83

RESULT 5

BM888450 519 bp mRNA linear EST 08-MAR-2002
 LOCUS TW108 Human Trabecular Meshwork cDNA library Homo sapiens cDNA
 DEFINITION clone 104447 5', mRNA sequence.

ACCESSION BM888450

VERSION BM888450.1 GI:19272194

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 519)

Wirtz,M.K., Samples,J.R., Xu,H., Severson,T. and Acott,T.S.

Expression Profile and Genome Location of cDNA Clones from an

Infant Human Trabecular Meshwork Library

Unpublished (2002)

JOURNAL

Contact: Wirtz MK

Glaucoma Genetics Lab

Oregon Health Sciences University

3375 S.W. Terwilliger Blvd., Portland, OR 97201-4197, USA

Tel: 503-494-4698

Fax: 503-494-6875

Email: wirtzm@ohsu.edu

Seq primer: T7 Reverse.

FEATURES

source

1..519
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="104447"
 /tissue type="eye"
 /cell type="trabecular meshwork"
 /dev stage="2 week to 2 year old infants"
 /lab_host="TPO10P"

/clone_lib="Human Trabecular Meshwork cDNA library"
 /note="Vector: pCDNA3; Site 1: EcoRI; Site 2: EcoRI; Human
 cDNA library made from mRNA isolated from trabecular
 meshwork cells established from eyes from 6 individuals,
 ages 2 weeks to 2 years. Cells were harvested at passages
 3 through 6. Invitrogen made a unidirectional cDNA library
 from the mRNA from the frozen cells using a pCDNA3 vector
 and TPO10P, host cells."

ORIGIN

Query Match 74.3%; Score 75; DB 5; Length 519;
 Best Local Similarity 94.0%; Pred. No. 7.5e-13;
 Matches 78; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 19 GTTTAAACCCGCTGATCAGCTCGACTGTGCTTCTAGTTGCCAGCCATCTGTTGTTGC 78
 |||

Db 333 GCTAGAGCTCGCTGATCAGCTCGACTGTGCTTCTAGTTGCCAGCCATCTGTTGTTGC 392

Qy 79 CCTCTCCCGCTGCTTCTCTTGAC 101
 |||

Db 393 CCTCTCCCGCTGCTTCTCTTGAC 415

RESULT 6

BM887817

LOCUS

DEFINITION BM887817 521 bp mRNA linear EST 08-MAR-2002
 TW553 Human Trabecular Meshwork cDNA library Homo sapiens cDNA
 clone 122060 5', mRNA sequence.

ACCESSION BM887817

VERSION BM887817.1 GI:19271561

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 521)

Wirtz,M.K., Samples,J.R., Xu,H., Severson,T. and Acott,T.S.

Expression Profile and Genome Location of cDNA Clones from an

Infant Human Trabecular Meshwork Library

Unpublished (2002)

JOURNAL

Contact: Wirtz MK

Glaucoma Genetics Lab

Oregon Health Sciences University

3375 S.W. Terwilliger Blvd., Portland, OR 97201-4197, USA

Tel: 503-494-4698

Fax: 503-494-6875

Email: wirtzm@ohsu.edu

Seq primer: T7 Reverse

High quality sequence stop: 350.

FEATURES

source

1..521
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="122060"
 /tissue type="eye"
 /cell type="trabecular meshwork"
 /dev stage="2 week to 2 year old infants"
 /lab_host="TPO10P"
 /clone_lib="Human Trabecular Meshwork cDNA library"
 /note="Vector: pCDNA3; Site 1: EcoRI; Site 2: EcoRI; Human
 cDNA library made from mRNA isolated from trabecular
 meshwork cells established from eyes from 6 individuals,
 ages 2 weeks to 2 years. Cells were harvested at passages
 3 through 6. Invitrogen made a unidirectional cDNA library
 from the mRNA from the frozen cells using a pCDNA3 vector
 and TPO10P, host cells."

ORIGIN

Query Match 74.3%; Score 75; DB 5; Length 521;
 Best Local Similarity 94.0%; Pred. No. 7.5e-13;
 Matches 78; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 19 GTTTAAACCCGCTGATCAGCTCGACTGTGCTTCTAGTTGCCAGCCATCTGTTGTTGC 78
 |||

Db 348 GCTAGAGCTCGCTGATCAGCTCGACTGTGCTTCTAGTTGCCAGCCATCTGTTGTTGC 407

Qy 79 CCTCTCCCGCTGCTTCTCTTGAC 101
 |||

Db 408 CCTCTCCCGCTGCTTCTCTTGAC 430

RESULT 7

CR154962/c

LOCUS

DEFINITION CR154962 233 bp DNA linear GSS 06-JUL-2004
 Forward strand read from insert in 3'HPRT insertion targeting and
 chromosome engineering clone MHP182j09, genomic survey sequence.

ACCESSION CR154962

VERSION CR154962.1 GI:49933807

```

KEYWORDS  GSS: genome survey sequence; MICER.
SOURCE     Mus musculus (house mouse)
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 233)
AUTHORS    Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
           Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
           Rogers,J. and Bradley,A.
TITLE      Direct Submission
JOURNAL    Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
           CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES   Location/Qualifiers
           source
             1..233
               /organism="Mus musculus"
               /mol_type="genomic DNA"
               /db_xref="taxon:10090"
               /clone="MHPP182j09"
               /clone_lib="MHPP"

ORIGIN
Query Match      73.3%; Score 74; DB 9; Length 233;
Best Local Similarity 88.9%; Pred. No. 1.3e-12;
Matches 80; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 12 AGGCGCGTTTAAACCGCTGATCAGCTCGACTGCTTCTAGTTCAGCCAGCCATCTGT 71
    |||
Db 223 ACGACCCCATGCATCGCGCTGATCAGCTCGACTGCTTCTAGTTCAGCCAGCCATCTGT 164
    |||

Qy 72 TGTGTGCCCCCTCCCGCTGCTTCTTGGAC 101
    |||
Db 163 TGTGTGCCCCCTCCCGCTGCTTCTTGGAC 134

RESULT 8
BM887701
LOCUS      BM887701
DEFINITION TM304 Human Trabecular Meshwork cDNA library Homo sapiens cDNA
           clone 107917 5', mRNA sequence.
ACCESSION  BM887701
VERSION    BM887701.1 GI:19271430
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 534)
AUTHORS    Wirtz,M.K., Samples,J.R., Xu,H., Severson,T. and Acott,T.S.
TITLE      Expression Profile and Genome Location of cDNA Clones from an
           Infant Human Trabecular Meshwork Library
JOURNAL    Unpublished (2002)
COMMENT    Contact: Wirtz MK
           Glaucoma Genetics Lab
           Oregon Health Sciences University
           3375 S.W. Terwilliger Blvd., Portland, OR 97201-4197, USA
           Tel: 503-494-4698
           Fax: 503-494-6875
           Email: wirtzm@ohsu.edu
           Seq primer: T7 Reverse.
FEATURES   Location/Qualifiers
           source
             1..534
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="107917"
               /tissue type="eye"
               /cell_type="trabecular meshwork"
               /dev_stage="2 week to 2 year old infants"
               /lab_host="TOP10F"
               /clone_lib="Human Trabecular Meshwork cDNA library"
               /note="Vector: pCDNA3; Site 1: EcoRI; Site 2: EcoRI; Human
               cDNA library made from mRNA isolated from trabecular
               meshwork cells established fom eyes from 6 individuals,

ages 2 weeks to 2 years. Cells were harvested at passages
3 through 6. Invitrogen made a unidirectional cDNA library
from the mRNA from the frozen cells using a pCDNA3 vector
and TPO10F", host cells."

ORIGIN
Query Match      72.7%; Score 73.4; DB 5; Length 534;
Best Local Similarity 92.8%; Pred. No. 2.4e-12;
Matches 77; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 19 GTTTAAACCGCTGATCAGCTCGACTGCTTCTAGTTCAGCCAGCCATCTGTGTTGC 78
    |||
Db 445 GCTAGAGCTCGCTGATCAGCTCGACTGCTTCTAGTTCAGCCAGCCATCTGTGTTGC 504
    |||

Qy 79 CCTCCCCCGTGCCTTCTTGGAC 101
    |||
Db 505 CCCTCCCCCGTGCCTTCTTGGAC 527

RESULT 9
CR074510/c
LOCUS      CR074510
DEFINITION Forward strand read from insert in 3'HPRT insertion targeting and
           chromosome engineering clone MHPP255d22, genomic survey sequence.
ACCESSION  CR074510
VERSION    CR074510.1 GI:49808100
KEYWORDS   GSS; genome survey sequence; MICER.
SOURCE     Mus musculus (house mouse)
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 132)
AUTHORS    Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
           Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
           Rogers,J. and Bradley,A.
TITLE      Direct Submission
JOURNAL    Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
           CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES   Location/Qualifiers
           source
             1..132
               /organism="Mus musculus"
               /mol_type="genomic DNA"
               /db_xref="taxon:10090"
               /clone="MHPP255d22"
               /clone_lib="MHPP"

ORIGIN
Query Match      72.5%; Score 73.2; DB 9; Length 132;
Best Local Similarity 90.7%; Pred. No. 2.2e-12;
Matches 78; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 12 AGGCGCGTTTAAACCGCTGATCAGCTCGACTGCTTCTAGTTCAGCCAGCCATCTGT 71
    |||
Db 93 ACGACCCCATGCATCGCGCTGATCAGCTCGACTGCTTCTAGTTCAGCCAGCCATCTGT 34
    |||

Qy 72 TGTGTGCCCCCTCCCGCTGCTTCTTCT 97
    |||
Db 33 TGTGTGCCCCCTCCCGCTGCTTCTTCT 8

RESULT 10
CG632479
LOCUS      CG632479
DEFINITION OST350781 Mus musculus 129Sv/Ev Mus musculus cDNA clone OST350781,
           mRNA sequence.
ACCESSION  CG632479
VERSION    CG632479.1 GI:37456328
KEYWORDS   GSS.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 329)

```

AUTHORS
 Zambrowicz,B.P., Abuin,A., Ramirez-Solis,R., Richter,L.J.,
 Piggott,J., Beltrandelio,H., Buxton,E.C., Edwards,J., Finch,R.A.,
 Friddle,C.J., Gupta,A., Hansen,G., Hu,Y., Huang,W., Jaing,C.,
 Key,B.W. Jr., Kipp,P., Kohlhauff,B., Ma,Z.-Q., Markesich,D.,
 Payne,R., Potter,D.G., Qian,N., Shaw,J., Schrick,J., Shi,Z.-Z.,
 Sparks,M.J., Van Slightenhorst,I., Vogel,P., Walke,W., Xu,N.,
 Zhu,Q., Person,C. and Sands,A.T.

TITLE
 Wnk1 kinase deficiency lowers blood pressure in mice: a gene-trap
 screen to identify potential targets for therapeutic intervention

JOURNAL
 Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)

COMMENT
 Contact: Zambrowicz BP
 OmniBank
 Lexicon Genetics Incorporated
 4000 Research Forest Drive, The Woodlands, TX 77381, USA
 Email: material@lexgen.com

**Gene trap sequence tag generated by 3' RACE from mouse ES cells as
 described in Zambrowicz et al (Nature. 1998 Apr 9;392(6676):608-11)**

Class: Gene Trap.

FEATURES
 source
 Location/Qualifiers
 1..329
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="129Sv/Ev"
 /db_xref="taxon:10090"
 /clone="OST350781"
 /cell_type="embryonic stem cell"
 /clone_lib="Mus musculus 129Sv/Ev"

ORIGIN
 Query Match 72.5%; Score 73.2; DB 9; Length 329;
 Best Local Similarity 96.2%; Pred. No. 2.5e-12;
 Matches 75; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

24 AACCGCTGATCAGCTCGACTGTGCTCTTAGTTGCCAGCATCTGTGTCCTC 83
 185 AGCTCGCTGATCAGCTCGACTGTGCTCTTAGTTGCCAGCATCTGTGTCCTC 244

84 CCCCCTGCTTCCTTGAC 101
 245 CCCCCTACCTTCCTTGAC 262

RESULT 11
 CR083191/c
 LOCUS
 DEFINITION Forward strand read from insert in 3'HPRT insertion targeting and
 chromosome engineering clone MHP263n05, genomic survey sequence.
 ACCESSION CR083191.1 GI:49816780
 VERSION GSS; genome survey sequence; MICER.
 KEYWORDS Mus musculus (house mouse)
 SOURCE Mus musculus
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 286)
 AUTHORS Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,I.,
 Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
 Rogers,J. and Bradley,A.
 TITLE Direct Submission
 JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. http://www.sanger.ac.uk/MICER
 FEATURES Location/Qualifiers
 source
 1..286
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /clone="MHP263n05"
 /clone_lib="MHP"

ORIGIN
 Query Match 71.5%; Score 72.2; DB 9; Length 286;
 Best Local Similarity 90.6%; Pred. No. 5.1e-12;
 Matches 77; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

12 AGGCGCCGTTTAAACCCGCTGATCAGCTCGACTGTGCTCTTAGTTGCCAGCATCTGT 71
 88 ACAGACCCCATCATCGCTCGCTGATCAGCTCGACTGTGCTCTTAGTTGCCAGCATCTGT 29

72 TGTGTGCCCCCTCCCGCTGCTTCC 96
 28 TGTGTGCCCCCTCCCGCTGCTTCC 4

RESULT 12
 BX988352/c
 LOCUS
 DEFINITION Forward strand read from insert in 3'HPRT insertion targeting and
 chromosome engineering clone MHP120m07, genomic survey sequence.
 ACCESSION BX988352
 VERSION 1 GI:49719810
 KEYWORDS GSS; genome survey sequence; MICER.
 SOURCE Mus musculus
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 130)
 AUTHORS Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,I.,
 Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
 Rogers,J. and Bradley,A.
 TITLE Direct Submission
 JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. http://www.sanger.ac.uk/MICER
 FEATURES Location/Qualifiers
 source
 1..130
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /clone="MHP120m07"
 /clone_lib="MHP"

ORIGIN
 Query Match 70.1%; Score 70.8; DB 9; Length 130;
 Best Local Similarity 86.7%; Pred. No. 1.2e-11;
 Matches 78; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

12 AGGCGCCGTTTAAACCCGCTGATCAGCTCGACTGTGCTCTTAGTTGCCAGCATCTGT 71
 120 ACAGACCCCATCATCGCTGATCAGCTCGACTGTGCTCTTAGTTGCCAGCATCTGT 61

72 TGTGTGCCCCCTCCCGCTGCTTCC 101
 60 TGTGTGCCCCCTCCCGCTGCTTCC 31

RESULT 13
 CR037248/c
 LOCUS
 DEFINITION Forward strand read from insert in 3'HPRT insertion targeting and
 chromosome engineering clone MHP31109, genomic survey sequence.
 ACCESSION CR037248
 VERSION 1 GI:49770303
 KEYWORDS GSS; genome survey sequence; MICER.
 SOURCE Mus musculus
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 75)
 AUTHORS Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,I.,
 Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
 Rogers,J. and Bradley,A.
 TITLE Direct Submission
 JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. http://www.sanger.ac.uk/MICER
 FEATURES Location/Qualifiers
 source
 1..75
 /organism="Mus musculus"

/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="MHPP31109"
/clone_lib="MHPP"

ORIGIN

Query Match 69.1%; Score 69.8; DB 9; Length 75;
Best Local Similarity 97.3%; Pred. No. 2.3e-11;
Matches 71; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 28 CGCTGATCAGCTCGACTGTGCTCTCTAGTTGCCAGCCATCTGTTTGGCCCTCCCCC 87
Db 74 CGCTGATCAGCTCGACTGTGCTCTCTAGTTGCCAGCCATCTGTTTGGCCCTCCCCC 15

Qy 88 GTGCCCTTCTTGA 100
Db 14 GTGCCCTTCTTGA 2

RESULT 14
BM887768 600 bp mRNA linear EST 08-MAR-2002
LOCUS TM397 Human Trabecular Meshwork cDNA library Homo sapiens cDNA
DEFINITION clone 119752 5', mRNA sequence.
ACCESSION BM887768
VERSION BM887768
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Query Match 68.5%; Score 69.2; DB 9; Length 166;
Best Local Similarity 85.6%; Pred. No. 4.1e-11;
Matches 77; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 12 AGGCCCGTTTAAACCGCTGATCAGCTCGACTGTGCTCTCTAGTTGCCAGCCATCTGT 71
Db 156 ACGACCCCATGCATCGCGCTGATCAGCTCGAGTGTCTTTCTAGTTGCCAGCCATCTGT 97

Qy 72 TGTTTGCCCTCCCGCTGCTTCTTGTAC 101
Db 96 TGTTTGCCCTCCCGCTGCTTCTTGTAC 67

Search completed: July 14, 2005, 23:23:04
Job time : 967.667 secs

/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="MHPP31109"
/clone_lib="MHPP"

ORIGIN

Query Match 69.1%; Score 69.8; DB 9; Length 75;
Best Local Similarity 97.3%; Pred. No. 2.3e-11;
Matches 71; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 28 CGCTGATCAGCTCGACTGTGCTCTCTAGTTGCCAGCCATCTGTTTGGCCCTCCCCC 87
Db 74 CGCTGATCAGCTCGACTGTGCTCTCTAGTTGCCAGCCATCTGTTTGGCCCTCCCCC 15

Qy 88 GTGCCCTTCTTGA 100
Db 14 GTGCCCTTCTTGA 2

RESULT 14
BM887768 600 bp mRNA linear EST 08-MAR-2002
LOCUS TM397 Human Trabecular Meshwork cDNA library Homo sapiens cDNA
DEFINITION clone 119752 5', mRNA sequence.
ACCESSION BM887768
VERSION BM887768
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Query Match 68.7%; Score 69.4; DB 5; Length 600;
Best Local Similarity 90.1%; Pred. No. 4.4e-11;
Matches 73; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 19 GTTTAAACCGCTGATCGACTGTGCTCTCTAGTTGCCAGCCATCTGTTTGTTC 78
Db 516 GCTAGAGCTCGCTGATCAGCTCGACTGTGCTCTCTAGTTGCCAGCCATCTGTTGTTC 575

ORIGIN

Query Match 68.7%; Score 69.4; DB 5; Length 600;
Best Local Similarity 90.1%; Pred. No. 4.4e-11;
Matches 73; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 19 GTTTAAACCGCTGATCGACTGTGCTCTCTAGTTGCCAGCCATCTGTTTGTTC 78
Db 516 GCTAGAGCTCGCTGATCAGCTCGACTGTGCTCTCTAGTTGCCAGCCATCTGTTGTTC 575

Result No.	Score	Query		DB	ID	Description
		Match	Length			
C 1	101	100.0	142	6	AR356490	Sequence
C 2	101	100.0	142	6	AR538046	Sequence
C 3	101	100.0	228	6	E00019	DNA coding
C 4	101	100.0	240	1	PM000DO	PM000DO
C 5	101	100.0	251	6	E00018	DNA coding
C 6	101	100.0	251	6	I01644	Sequence 1
C 7	101	100.0	344	11	HUMUT5345	Human chromosome 1
C 8	101	100.0	400	6	BD195256	Nucleotide
C 9	101	100.0	456	6	E00892	Synthetic D
C 10	101	100.0	456	6	E00156	DNA fragment
C 11	101	100.0	456	6	E01274	DNA encoding
C 12	101	100.0	456	6	E01302	DNA encoding
C 13	101	100.0	466	6	AX260098	Sequence
C 14	101	100.0	573	6	AX260150	Sequence
C 15	101	100.0	693	6	A43586	Sequence 1
C 16	101	100.0	693	6	AR116755	Sequence
C 17	101	100.0	998	1	AY559171	Pseudomon
C 18	101	100.0	1011	1	SMTEA0GE	X97254 S.marcesc
C 19	101	100.0	1012	2	CEC11F0G	Z92776 Caenorhabdi

ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 142)
AUTHORS Kunsch,C.A., Choi,G.A., Barash,S.C., Dillon,P.J., Fannon,M.R. and Rosen,C.A.
TITLE Staphylococcus aureus polynucleotides and sequences
JOURNAL Patent: US 6737248-A 2608 18-MAY-2004;
FEATURES Location/Qualifiers
source 1..142
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 142;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGAATGTTAGAAAAATAACAATAG 60
Db 107 AGGGTTATTGCTCATGAGCGGATACATATTGAATGTTAGAAAAATAACAATAG 48
Qy 61 GGGTTCGGCGACATTTCCCGGAAAAGTGCCACCTGACGTC 101
Db 47 GGGTTCGGCGACATTTCCCGGAAAAGTGCCACCTGACGTC 7
RESULT 3
E00019/c
LOCUS DNA coding for Escherichia coli penicillinase.
DEFINITION E00019
ACCESSION E00019
VERSION E00019.1 GI:2168327
KEYWORDS JP 1981154999-A/2.
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE 1 (bases 1 to 228)
AUTHORS Uotutaa,G. and Karen,T.
TITLE SYNTHESIS OF SELECTIVE AGED PROTEIN OR POLYPEPTIDE IN BACTERIA
JOURNAL Patent: JP 1981154999-A 2 30-NOV-1981;
COMMENT UNIV HARVARD
OS Escherichia coli
PN JP 1981154999-A/2
PD 30-NOV-1981
PF 09-APR-1981 JP 1981052488
PI 11-APR-1980 US 80 139225
PI UORUTAA GIRUBATO, KAREN TARUMATSUJI
PC C12P21/00,C07H21/00,C12N1/00,C12N15/00//C12R1/19; CC
strandedness: Double;
CC topology: Linear;
CC anti-sense: No;
CC *source: clone=pKT218;
FH Key Location/Qualifiers
FH CDS 210..>228
FT /product='E.coli penicillinase'.
FT Location/Qualifiers
FEATURES
source 1..228
/organism="Escherichia coli"
/mol_type="genomic DNA"
/db_xref="taxon:562"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 228;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGAATGTTAGAAAAATAACAATAG 60
Db 175 AGGGTTATTGCTCATGAGCGGATACATATTGAATGTTAGAAAAATAACAATAG 116
Qy 61 GGGTTCGGCGACATTTCCCGGAAAAGTGCCACCTGACGTC 101

Db 115 GGGTTCGGCGACATTTCCCGGAAAAGTGCCACCTGACGTC 75
RESULT 4
PMOENDO/c
LOCUS DNA 240 bp linear BCT 26-APR-1993
DEFINITION Plasmid pMM110 region of endo VII cleavage sites near cruciform structures.
ACCESSION M10199
VERSION M10199.1 GI:150826
KEYWORDS
SOURCE Plasmid pMM110
ORGANISM Plasmid pMM110
other sequences; plasmids.
REFERENCE 1 (bases 1 to 240)
AUTHORS Kemper,B., Jensch,P., von Depka-Prondzynski,M., Fritz,H.J., Borgmeyer,U. and Mizuuchi,K.
TITLE Resolution of Holliday structures by endonuclease VII as observed in interactions with cruciform DNA
JOURNAL Cold Spring Harb. Symp. Quant. Biol. 49, 815-825 (1984)
MEDLINE 85153063
PUBMED 6397324
COMMENT Original source text: Plasmid pMM110 DNA.
FEATURES Location/Qualifiers
source 1..240
/organism="Plasmid pMM110"
/mol_type="genomic DNA"
/db_xref="taxon:2599"
/plasmid="Plasmid pMM110"
ORIGIN Unreported.
Query Match 100.0%; Score 101; DB 1; Length 240;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGAATGTTAGAAAAATAACAATAG 60
Db 151 AGGGTTATTGCTCATGAGCGGATACATATTGAATGTTAGAAAAATAACAATAG 92
Qy 61 GGGTTCGGCGACATTTCCCGGAAAAGTGCCACCTGACGTC 101
Db 91 GGGTTCGGCGACATTTCCCGGAAAAGTGCCACCTGACGTC 51
RESULT 5
E00018/c
LOCUS DNA coding for Escherichia coli penicillinase.
DEFINITION E00018
ACCESSION E00018
VERSION E00018.1 GI:2168326
KEYWORDS JP 1981154999-A/1.
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE 1 (bases 1 to 251)
AUTHORS Uotutaa,G. and Karen,T.
TITLE SYNTHESIS OF SELECTIVE AGED PROTEIN OR POLYPEPTIDE IN BACTERIA
JOURNAL Patent: JP 1981154999-A 1 30-NOV-1981;
COMMENT UNIV HARVARD
OS Escherichia coli
PN JP 1981154999-A/1
PD 30-NOV-1981
PF 09-APR-1981 JP 1981052488
PI 11-APR-1980 US 80 139225
PI UORUTAA GIRUBATO, KAREN TARUMATSUJI
PC C12P21/00,C07H21/00,C12N1/00,C12N15/00//C12R1/19; CC
strandedness: Double;
CC topology: Linear;
CC anti-sense: No;
CC fragment type: N-Terminal Fragment;
CC *source: clone=pKT241;


```

FH Key Location/Qualifiers
FH CDS 210..>252
FT /product='E.coli penicillinase' FT
TATA_signal 190..196
FEATURES             Location/Qualifiers
     source          1..251
                     /organism='Escherichia coli'
                     /mol_type='genomic DNA'
                     /db_xref='taxon:562'
ORIGIN
Query Match          100.0%; Score 101; DB 6; Length 251;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTATTAGAAAATAAACAAATAG 60
    |||||||
Db 175 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTATTAGAAAATAAACAAATAG 116
    |||||||
QY 61 GGGTTCGGCGCACATTTCCCGAAAAAGTCCACCTGACGTC 101
    |||||||
Db 115 GGGTTCGGCGCACATTTCCCGAAAAAGTCCACCTGACGTC 75

RESULT 6
I01644/c
LOCUS                251 bp ss-DNA linear PAT 18-MAY-1993
DEFINITION           Sequence 1 from Patent US 4338397.
ACCESSION             I01644
VERSION               I01644.1 GI:267685
KEYWORDS              .
SOURCE               Unknown.
ORGANISM              Unclassified.
REFERENCE             1 (bases 1 to 251)
AUTHORS              Gilbert,W. and Talmadge,K.
TITLE                Mature protein synthesis
JOURNAL              Patent: US 4338397-A 1 06-JUL-1982;
                    President and Fellows of Harvard College; Cambridge, MA
FEATURES             Location/Qualifiers
     source          1..251
                     /organism='unknown'
                     /mol_type='unassigned DNA'
ORIGIN

Query Match          100.0%; Score 101; DB 6; Length 251;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTATTAGAAAATAAACAAATAG 60
    |||||||
Db 175 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTATTAGAAAATAAACAAATAG 116
    |||||||
QY 61 GGGTTCGGCGCACATTTCCCGAAAAAGTCCACCTGACGTC 101
    |||||||
Db 115 GGGTTCGGCGCACATTTCCCGAAAAAGTCCACCTGACGTC 75

RESULT 7
HUMUT5345
LOCUS                344 bp DNA linear STS 26-JUL-1993
DEFINITION           Human chromosome 8 STS UT5345, sequence tagged site.
ACCESSION             L18624
VERSION               L18624.1 GI:308338
KEYWORDS              STS; PCR primer; STS sequence; microsatellite marker;
                    microsatellite repeat; repeat polymorphism; sequence tagged site.
SOURCE               Homo sapiens
ORGANISM              Homo sapiens
REFERENCE             1 (bases 1 to 344)
AUTHORS              Gerken,S.C., Matsunami,N., Lawrence,E., Carlson,M., Moore,M.,
Ballard,L., Melis,R., Robertson,M., Bradley,P., Elsner,T.,
Tingey,A., Rodriguez,P., Albertsen,H., Lalouel,J.-M. and White,R.
Genetic and physical mapping of simple sequence repeat containing
sequence tagged sites from the human genome
Unpublished (1993)
Original source text: Homo sapiens DNA.
Submitted by: Utah Center for Human Genome Research University of
Utah, Dept. of Human Genetics
2160 Eccles Institute of Human Genetics
Salt Lake City, UT 84112
e-mail: sts@corona.med.utah.edu
Primer A: GAGCAAAAACAGAGGCAAAATGC
Primer B: TTCGGGAAATGTCCCGAACC
32P-label: B Primer
PCR Profile:
Initial Denaturation: 94C 300sec
PCR Cycles: 30
Denaturation: 94C 10sec
Annealing: 60C 10sec
Extension: 72C 20sec
Mg++: 2mM
Gel: Acrylamide 7%, Formamide 32%, Urea 34%
Alleles: 2.
FEATURES             Location/Qualifiers
     source          1..344
                     /organism='Homo sapiens'
                     /mol_type='genomic DNA'
                     /db_xref='taxon:9606'
                     /map='8'
                     36..224
                     /standard_name='STS UT5345'
                     36..60
                     complement(202..224)
     primer_bind
     primer_bind
ORIGIN

Query Match          100.0%; Score 101; DB 11; Length 344;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTATTAGAAAATAAACAAATAG 60
    |||||||
Db 141 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTATTAGAAAATAAACAAATAG 200
    |||||||
QY 61 GGGTTCGGCGCACATTTCCCGAAAAAGTCCACCTGACGTC 101
    |||||||
Db 201 GGGTTCGGCGCACATTTCCCGAAAAAGTCCACCTGACGTC 241

RESULT 8
BD195256/c
LOCUS                400 bp DNA linear PAT 17-JUL-2003
DEFINITION           Nucleotide sequence of Escherichia coli pathogenicity islands.
ACCESSION             BD195256
VERSION               BD195256.1 GI:33005021
KEYWORDS              JP 2002513277-A/43.
SOURCE               unidentified
ORGANISM              unclassified.
REFERENCE             1 (bases 1 to 400)
AUTHORS              Dillon,P.J., Choi,G.H. and Welch,R.A.
TITLE                Nucleotide sequence of Escherichia coli pathogenicity islands
JOURNAL              Patent: JP 2002513277-A 43 08-MAY-2002;
                    HUMAN GENOME SCIENCES INC./WISCONSIN ALUMNI RESEARCH FOUNDATION
COMMENT              OS Unidentified
                    PN JP 2002513277-A/43
                    PD 08-MAY-2002
                    PF 21-NOV-1997 JP 1998523916
                    PR 22-NOV-1996 US 60/031626,14-OCT-1997 US 60/061953 PI
                    PC C12N15/11,C12N15/63,C07K16/12,G01N33/569,G06F17/30,G11B7/00 CC
                    Strandedness: Double;
                    CC Topology: Linear;
                    CC Nucleotide sequence of Escherichia coli pathogenicity islands

```

```

FH Key Location/Qualifiers
FT source 1..400
FT /organism='Unidentified'.

FEATURES
    source
        Location/Qualifiers
            1..400
            /organism='unidentified'
            /mol_type='genomic DNA'
            /db_xref='taxon:32644'

ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 400;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGTTATTGCTCATGAGCGGATACATATTGTAATGATTTAGAAAATAAACAATAG 60
    |||||||
Db 165 AGGTTATTGCTCATGAGCGGATACATATTGTAATGATTTAGAAAATAAACAATAG 106
    |||||||

Qy 61 GGGTTCGGCGACATTTCCCGGAAAGTGCCACCTGACGTC 101
    |||||||
Db 105 GGGTTCGGCGACATTTCCCGGAAAGTGCCACCTGACGTC 65
    |||||||

RESULT 9
E00892/c
LOCUS E00892 456 bp DNA linear PAT 29-SEP-1997
DEFINITION Synthetic DNA encoding fused polypeptide between E coli
beta-lactamase and human beta-urogastrone.
ACCESSION E00892
VERSION E00892.1 GI:2169153
KEYWORDS JP 1986149089-A/1.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 456)
AUTHORS Okai,H., Momota,Y., Kumakura,T., Tochifusa,N., Kitazawa,T.,
Ojida,K. and Matsushiro,A.
TITLE POLYPEPTIDE SECRETION DEVELOPMENT VECTOR, MICROORGANISM TRANSFORMED
WITH SAME, AND PRODUCTION OF POLYPEPTIDE WITH MICROORGANISM
JOURNAL Patent: JP 1986149089-A 1 07-JUL-1986;
EARTH CHEM CORP LTD
COMMENT OS Artificial gene
OC Artificial sequence; Genes.
PN JP 1986149089-A/1
PD 07-JUL-1986
PF 21-DEC-1984 JP 1984271206
PI OKAI HIDEO, MOMOTA YUTAKA, KUMAKURA TAKESHI,
PI TOCHIFUSA NORIYUKI,
PI KITAZAWA TOSHIKI, OJIDA KAZUHIRO, MATSUSHIRO AIZO PC
C12N15/00,C12N1/20,C12P21/00,(C12N1/20,C12R1:19),(C12P21/00,PC
C12R1:19);
CC strandedness: Double;
CC topology: linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: strain=Hb101;
CC *source: clone=pvG201;
CC Feature is identified by experimental;
FH Key Location/Qualifiers
FH promoter 125..170
FT of beta-lactamase
FT RBS 200..203
FT CDS 209..438
FT /product='beta-urogastrone precursor' FT
FT sig_peptide 209..277
FT /product='signal peptide of beta-lactonase' FT
FT mat_peptide 278..435
FT /product='beta-urogastrone mature peptide'.
FT Location/Qualifiers
    1..456
    /organism='synthetic construct'
    /mol_type='genomic DNA'

FEATURES
    source
        Location/Qualifiers
            1..456
            /organism='synthetic construct'
            /mol_type='genomic DNA'

```

```

ORIGIN
/db_xref='taxon:32630'

Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGTTATTGCTCATGAGCGGATACATATTGTAATGATTTAGAAAATAAACAATAG 60
    |||||||
Db 173 AGGTTATTGCTCATGAGCGGATACATATTGTAATGATTTAGAAAATAAACAATAG 114
    |||||||

Qy 61 GGGTTCGGCGACATTTCCCGGAAAGTGCCACCTGACGTC 101
    |||||||
Db 113 GGGTTCGGCGACATTTCCCGGAAAGTGCCACCTGACGTC 73
    |||||||

RESULT 10
E01156/c
LOCUS E01156 456 bp DNA linear PAT 29-SEP-1997
DEFINITION DNA fragment which secretes beta urogastrone.
ACCESSION E01156
VERSION E01156.1 GI:2169415
KEYWORDS JP 1987083890-A/1.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 456)
AUTHORS Yoshikawa,K., Momota,Y., Kajifusa,N., Koide,T. and Okai,H.
TITLE POLYPEPTIDE SECRETION EXPRESSION VECTOR, MICROORGANISM TRANSFORMED
BY SAID VECTOR AND PRODUCTION OF POLYPEPTIDE USING SAID
JOURNAL Patent: JP 1987083890-A 1 17-APR-1987;
EARTH CHEM CORP LTD
COMMENT OS Artificial gene
OC Artificial sequence; Genes.
PN JP 1987083890-A/1
PD 17-APR-1987
PF 09-OCT-1985 JP 1985225393
PI YOSHIKAWA KAZUTOSHI, MOMOTA YUTAKA, KAJIFUSA NORIYUKI, PI
KOIDE TAKAO,
PI OKAI HIDEO
PC C12N15/00,C12N1/20,C12P21/00,(C12N1/20,C12R1:19),(C12N1/20,PC
C12R1:125);
CC strandedness: Double;
CC topology: linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: clone=pvG201;
FH Key Location/Qualifiers
FH promoter 125..170
FT /note='beta lactamase promoter' FT RBS
FT CDS 209..439
FT /product='beta urogastrone'
FT sig_peptide 209..277
FT mat_peptide 278..436
FT /product='beta urogastrone'.
FH Key Location/Qualifiers
    1..456
    /organism='synthetic construct'
    /mol_type='genomic DNA'
    /db_xref='taxon:32630'

ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGTTATTGCTCATGAGCGGATACATATTGTAATGATTTAGAAAATAAACAATAG 60
    |||||||
Db 173 AGGTTATTGCTCATGAGCGGATACATATTGTAATGATTTAGAAAATAAACAATAG 114
    |||||||

Qy 61 GGGTTCGGCGACATTTCCCGGAAAGTGCCACCTGACGTC 101
    |||||||

```

```

Db      113 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 73
|||||
RESULT 11
E01274/c
LOCUS   E01274
DEFINITION DNA encoding beta-urogastron fused with DNA encoding a promoter and signal peptide of beta-lactamase.
ACCESSION E01274
VERSION   E01274.1 GI:2169533
KEYWORDS  JP 1987179398-A/1.
SOURCE    synthetic construct
ORGANISM  other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 456)
AUTHORS   Okai,H., Kumakura,T., Kawamoto,S., Adachi,S., Matsubara,A.,
           Ojida,K., Yano,M., Mihara,S., Matsuhiro,A. and Yanaihara,N.
TITLE     PRODUCTION OF BETA-UROGASTRONE
JOURNAL   Patent: JP 1987179398-A 1 06-AUG-1987;
           EARTH CHEM CORP LTD
COMMENT   OS Artificial gene
           OS Artificial sequence; Genes.
           OS Homo sapiens
           PN JP 1987179398-A/1
           PD 06-AUG-1987
           PF 31-JAN-1986 JP 1986021032
           PI OKAI HIDEO, KUMAKURA TAKESHI, KAWAMOTO SHOICHI, ADACHI SHOJI,
           PI MATSUBARA AKIMASA, OJIDA KAZUHIRO, YANO MAKI, MIHARA SHIGERU,
           PI MATSUHISHIRO AIZO, YANAIHARA NOBORU
           PC C12P21/00,C12N15/00,(C12P21/00,C12R1:91);
           CC strandedness: Double;
           CC topology: Linear;
           CC hypothetical: No;
           CC anti-sense: No;
           FH Key
           FH Location/Qualifiers
           FT Promoter 125..170
           FT RBS 200..203
           FT sig_peptide 209..277
           FT mat_peptide 278..436
           FT /product='beta-urogastron'
           FT CDS 209..439
           FT /product='beta-urogastron'.
FEATURES
source
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20; Mismatches 0; Indels 0; Gaps 0;
Matches 101; Conservative 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTTATTAGAAAAATAAACAAATAG 60
Db 173 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTTATTAGAAAAATAAACAAATAG 114
|||||
Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101
Db 113 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 73
|||||
RESULT 13
AX260098/c
LOCUS   AX260098
DEFINITION Sequence 60 from Patent WO0172774.
ACCESSION AX260098
VERSION   AX260098.1 GI:16509129
KEYWORDS
SOURCE    Drosophila melanogaster (fruit fly)
ORGANISM  Drosophila melanogaster
           Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
           Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
           Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1
AUTHORS   Deak,P., Glover,D.M. and Midgley,C.
TITLE     Cell cycle progression proteins
JOURNAL   Patent: WO 0172774-A 60 04-OCT-2001;
           Cyclacel Limited (GB)
FEATURES
source
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20; Mismatches 0; Indels 0; Gaps 0;
Matches 101; Conservative 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTTATTAGAAAAATAAACAAATAG 60
Db 173 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTTATTAGAAAAATAAACAAATAG 114
|||||
Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101
Db 113 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 73
|||||
RESULT 12
E01302/c
LOCUS   E01302
DEFINITION DNA encoding human beta-urogastrone fused with DNA encoding promoter and signal peptide of beta-lactamase.
ACCESSION E01302
VERSION   E01302.1 GI:2169561
KEYWORDS  JP 1987190083-A/1.
SOURCE    synthetic construct
ORGANISM  other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 456)
AUTHORS   Okai,H., Kumakura,T., Kawamoto,S., Adachi,S., Matsubara,A.,
           Ojida,K., Yano,M., Mihara,S., Matsuhiro,A. and Yanaihara,N.
TITLE     PRODUCTION OF BETA-UROGASTRONE
JOURNAL   Patent: JP 1987190083-A 1 20-AUG-1987;
           EARTH CHEM CORP LTD
COMMENT   OS Artificial gene
           OS Artificial sequence; Genes.
           OS Homo sapiens
           PN JP 1987190083-A/1
           PD 20-AUG-1987
           PF 14-FEB-1986 JP 1986031415
           PI OKAI HIDEO, KUMAKURA TAKESHI, KAWAMOTO SHOICHI, KOIDE TAKAO,
           PI MOMOTA YUTAKA
           PC C12N15/00,C07H21/04,C12N1/00,C12P21/02,(C12N1/00,C12R1:19), PC
           (C12P21/02);
           CC strandedness: Double;
           CC topology: Linear;
           CC hypothetical: No;
           CC anti-sense: No;
           FH Key
           FH Location/Qualifiers
           FT Promoter 125..170
           FT RBS 200..203
           FT sig_peptide 209..277
           FT mat_peptide 278..436
           FT /product='human beta-urogastrone' FT CDS
           FT 209..439
           FT /product='human beta-urogastrone'.
FEATURES
source
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20; Mismatches 0; Indels 0; Gaps 0;
Matches 101; Conservative 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTTATTAGAAAAATAAACAAATAG 60
Db 173 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTTATTAGAAAAATAAACAAATAG 114
|||||
Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101
Db 113 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 73
|||||
RESULT 10
AX260098/c
LOCUS   AX260098
DEFINITION Sequence 60 from Patent WO0172774.
ACCESSION AX260098
VERSION   AX260098.1 GI:16509129
KEYWORDS
SOURCE    Drosophila melanogaster (fruit fly)
ORGANISM  Drosophila melanogaster
           Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
           Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
           Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1
AUTHORS   Deak,P., Glover,D.M. and Midgley,C.
TITLE     Cell cycle progression proteins
JOURNAL   Patent: WO 0172774-A 60 04-OCT-2001;
           Cyclacel Limited (GB)
FEATURES
source
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20; Mismatches 0; Indels 0; Gaps 0;
Matches 101; Conservative 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTTATTAGAAAAATAAACAAATAG 60
Db 173 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTTATTAGAAAAATAAACAAATAG 114
|||||
Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101
Db 113 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 73
|||||

```

/db_xref="taxon:7227"

ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 466;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGATGTATTAGAAAAATAAACAAATAG 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
280 AGGGTTATTGCTCATGAGCGGATACATATTGATGTATTAGAAAAATAAACAAATAG 221
Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
220 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 180

RESULT 14

AX260150/c AX260150 573 bp DNA linear PAT 26-OCT-2001
LOCUS
DEFINITION Sequence 112 from Patent WO0172774.
ACCESSION AX260150
VERSION AX260150.1 GI:16509172

KEYWORDS

SOURCE

ORGANISM

Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE

AUTHORS

TITLE

JOURNAL

Cyclacel

Limited (GB)

Patent: WO 01/72774-A 112 04-OCT-2001;

Cyclacel

Limited (GB)

Patent: WO 01/72774-A 112 04-OCT-2001;

Cyclacel

Limited (GB)

Patent: WO 01/72774-A 112 04-OCT-2001;

Cyclacel

Limited (GB)

Patent: WO 01/72774-A 112 04-OCT-2001;

Cyclacel

Limited (GB)

Patent: WO 01/72774-A 112 04-OCT-2001;

Cyclacel

Limited (GB)

Patent: WO 01/72774-A 112 04-OCT-2001;

Cyclacel

Limited (GB)

Patent: WO 01/72774-A 112 04-OCT-2001;

Cyclacel

Limited (GB)

Patent: WO 01/72774-A 112 04-OCT-2001;

Cyclacel

Limited (GB)

Patent: WO 01/72774-A 112 04-OCT-2001;

Cyclacel

Limited (GB)

Patent: WO 01/72774-A 112 04-OCT-2001;

Cyclacel

Limited (GB)

Patent: WO 01/72774-A 112 04-OCT-2001;

Cyclacel

Limited (GB)

Patent: WO 01/72774-A 112 04-OCT-2001;

Cyclacel

Limited (GB)

Patent: WO 01/72774-A 112 04-OCT-2001;

Cyclacel

Limited (GB)

Patent: WO 01/72774-A 112 04-OCT-2001;

Cyclacel

Limited (GB)

Other publication AU 7615494 950327.

FEATURES

source

Location/Qualifiers

1..693

/organism="Cuphea lanceolata"

/mol_type="unassigned DNA"

/db_xref="taxon:3930"

/clone="CLKASIG8"

/clone_lib="Genomic Lambda Fix II"

ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 693;
Best Local Similarity 100.0%; Pred. No. 8.8e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGATGTATTAGAAAAATAAACAAATAG 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
592 AGGGTTATTGCTCATGAGCGGATACATATTGATGTATTAGAAAAATAAACAAATAG 651
Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
652 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 692

Search completed: July 14, 2005, 14:03:32

Job time : 756.618 secs

ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 573;
Best Local Similarity 100.0%; Pred. No. 8.8e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGATGTATTAGAAAAATAAACAAATAG 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
355 AGGGTTATTGCTCATGAGCGGATACATATTGATGTATTAGAAAAATAAACAAATAG 296
Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
295 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 255

RESULT 15

A43586

LOCUS

A43586

DEFINITION

Sequence 11 from Patent WO9507357.

ACCESSION

A43586

VERSION

A43586.1

KEYWORDS

Cuphea lanceolata

Cuphea lanceolata

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; Myrtales; Lythraceae; Cuphea.

1 (bases 1 to 693)

Toepfer, R., Bautor, J., Bothmann, H., Filsak, E.,

Hoerhke-Grandpierre, C., Klein, B., Martini, N., Mueller, A.,

Schulte, W., Voetz, M., Walek, J. and Scheil, J.

PROMOTERS

Patent: WO 9507357-A 11 16-MAR-1995;

MAX PLANCK GESELLSCHAFT (DE)

Other publication CA 2169093 950316

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:35:42 ; Search time 142.398 Seconds
(without alignments)
4198.742 Million cell updates/sec

Title: US-09-482-682-44_COPY_7860_7960
Perfect score: 101
Sequence: 1 aggtgtattgtctcatgagc.....gaaagtgcacctgacgtc 101

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues.

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	101	100.0	142	2	AAV76919
C 2	101	100.0	228	1	Aan10032 Sequence
C 3	101	100.0	251	1	Aan10031 Sequence
C 4	101	100.0	400	2	AAV31229 E. coli J
C 5	101	100.0	456	1	Aan60824 Plasmid p
C 6	101	100.0	456	1	Aan71080 Sequence
C 7	101	100.0	456	1	Aan70833 Beta-urog
C 8	101	100.0	456	1	Aan81765 Sequence
C 9	101	100.0	466	6	ABA90413 Drosophil
C 10	101	100.0	487	2	AAx21173 Polynucle
C 11	101	100.0	535	2	AAx21149 Polynucle
C 12	101	100.0	573	6	ABA90456 Drosophil
C 13	101	100.0	605	12	ADH58311 Electroph
C 14	101	100.0	776	4	AAx30560 DNA encod
C 15	101	100.0	776	4	AAx27819 DNA encod
C 16	101	100.0	776	4	ABK42984 Genomic s
C 17	101	100.0	776	4	AA107344 Human rep
C 18	101	100.0	776	4	AA103229 Human rep
C 19	101	100.0	776	4	AA106588 Human rep
C 20	101	100.0	776	4	AA107340 Human rep

C 21	101	100.0	776	5	ABA14573
C 22	101	100.0	776	5	AAS34681
C 23	101	100.0	776	8	ADA41574
C 24	101	100.0	776	8	ACC50905
C 25	101	100.0	776	8	ABZ71508
C 26	101	100.0	776	9	ADB91869
C 27	101	100.0	776	9	ADB61140
C 28	101	100.0	776	10	ADB94622
C 29	101	100.0	776	10	ADC74663
C 30	101	100.0	776	10	ADA57709
C 31	101	100.0	776	12	ADN41551
C 32	101	100.0	845	4	AAS30559
C 33	101	100.0	845	4	AAS27818
C 34	101	100.0	845	4	ABK42983
C 35	101	100.0	845	4	AAS41807
C 36	101	100.0	845	4	AAS41855
C 37	101	100.0	845	4	AAK85485
C 38	101	100.0	845	4	AAK85434
C 39	101	100.0	845	4	AA107343
C 40	101	100.0	845	4	AA106587
C 41	101	100.0	845	4	AA107339
C 42	101	100.0	845	4	AA103228
C 43	101	100.0	845	5	ABA14572
C 44	101	100.0	845	5	AAS34680
C 45	101	100.0	845	9	ADB61139

ALIGNMENTS

RESULT 1
AAV76919/c
ID AAV76919 standard; DNA; 142 BP.

XX AAV76919;

DT 16-MAR-1999 (first entry)

DE Staphylococcus aureus contig SEQ ID #2608.

XX Computer readable medium; vaccine; S.aureus infection; immunodetection;
KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW skin infection; surgical wound infection; scalded skin syndrome;
KW toxic shock syndrome; ds.

XX Staphylococcus aureus.

XX EP786519-A2.

XX 30-JUL-1997.

XX 07-JAN-1997; 97EP-00100117.

XX 05-JAN-1996; 96US-0009861P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Kunsch CA, Choi GH, Barash SC, Dillon PJ, Fannon MR, Rosen CA;

XX WPI; 1997-374922/35.

XX Polynucleotide(s) and proteins derived from Staphylococcus aureus -
stored on computer readable medium and used in the production of anti-
S.aureus vaccines.

PS Claim 1; Page 2287; 3271pp; English.

XX This sequence represents one of 5191 Staphylococcus aureus DNA sequences
of the invention. The DNA sequences are recorded on a computer readable
medium, preferably selected from a floppy or hard disk, random access
memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
the S.aureus DNA sequences allows putative functions to be assigned so
that protein-encoding or regulatory regions of commercial, therapeutic or

CC 3' end of pK218 was attached to the signal DNA sequence of the DNA
CC fragment (CB6) for rat preproinsulin (see AAN10034)

XX SQ Sequence 251 BP; 74 A; 45 C; 50 G; 82 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 1; Length 251;
Best Local Similarity 100.0%; Pred. No. 2.3e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTAGAAAAATAAACAATAG 60
DB |||||||
175 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTAGAAAAATAAACAATAG 116
QY 61 GGGTTCCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 101
DB |||||||
115 GGGTTCCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 75

RESULT 4

AAV31229/c
ID AAV31229 standard; DNA; 400 BP.

XX AC AAV31229;

DT 01-OCT-1998 (first entry)

XX E. coli J96 pathogenicity island contig #43.

XX PAI; pathogenicity island; uropathogenic E. coli detection; PAI IV; pheR;
KW PAI V; pheV; vaccine; protective immune response; ds.

XX Escherichia coli.

XX WO9822575-A2.

XX 28-MAY-1998.

XX 21-NOV-1997; 97WO-US021347.

XX 22-NOV-1996; 96US-0031626P.

PR 14-OCT-1997; 97US-0061953P.

XX (HUMA-) HUMAN GENOME SCI INC.

PA (UYWI-) UNIV WISCONSIN.

XX Dillon PJ, Choi GH, Welch RA;

XX WPI; 1998-312461/27.

PT New isolated uropathogenic E. coli nucleotide sequences - used to develop
PT products for the detection of pathogenic E. coli and to elicit an immune
PT response to pathogenic E. coli.

PS Claim 21; Page 140-141; 250pp; English.

XX This sequence represents a E. coli strain J96 contig containing
CC pathogenicity island (PAI) sequences, and represents a nucleic acid
CC molecule of the invention. PAIs are large fragments of DNA which comprise
CC pathogenicity determinants. The sequences of the invention are taken from
CC PAI IV and PAI V. PAI IV is located at approximately 64 min (near pheV)
CC on the E. coli chromosome and is greater than 170 kb. PAI V is located at
CC approximately 94 min (at pheR) on the E. coli chromosome and is
CC approximately 160 kb in size. Antibodies specific to the proteins encoded
CC by the PAI open reading frames of the invention can be used in kits to
CC detect uropathogenic E. coli. The proteins are used in vaccines to elicit
CC a protective immune response in an animal to the uropathogenic E. coli
CC strain J96

XX SQ Sequence 400 BP; 106 A; 77 C; 91 G; 126 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 2; Length 400;

Best Local Similarity 100.0%; Pred. No. 2.5e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTAGAAAAATAAACAATAG 60
DB |||||||
165 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTAGAAAAATAAACAATAG 106

QY 61 GGGTTCCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 101
DB |||||||
105 GGGTTCCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 65

RESULT 5

AAN60624/c
ID AAN60624 standard; DNA; 456 BP.

XX AC AAN60624;

DT 25-MAR-2003 (revised)

DT 29-OCT-1991 (first entry)

XX Plasmid pUG201 sequence encoding beta-urogastrone.

XX Beta-lactamase signal peptide; pGH54; pGH55; ss.

XX Synthetic.

XX Key Location/Qualifiers

FT promoter 125..170

FT /*tag= a

FT RBS 200..203

FT /*tag= b

FT CDS 209..439

FT /*tag= c

FT sig_peptide 209..277

FT /*tag= d

FT /label= Beta-lactamase signal peptide

FT mat_peptide 278..436

FT /*tag= e

FT /label= Beta-urogastrone

XX WO8603779-A.

XX 03-JUL-1986.

XX 19-DEC-1985; 85WO-JP000696.

XX 21-DEC-1984; 84JP-00271206.

XX (EART) EARTH CHEM CO LTD.

PA (OHGA/) OHGAI H.

XX Ohgai H, Momota H, Kuwakura T, Kajifusa N, Kitazawa T, Oshiden K;

XX WPI; 1986-182911/28.

XX P-PSDB; AAP60678.

XX Recombinant vector for polypeptide secretion - contains signal peptide
PT sequence directly bonded to peptide-coding sequence.

XX Disclosure; Table 4; 79pp; Japanese.

XX The plasmid produces secreted beta-urogastrone in a transformed
CC expression system. Similar plasmids may be constructed where the
CC secretion signal may be coupled with eg. somatostatin, insulin, growth
CC hormone, interferon, IL-2, gastric inhibitory peptide, influenza B SA,
CC epidermal growth factor and thymosine-beta4. (Updated on 25-MAR-2003 to
CC correct PA field.)

XX SQ Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 1; Length 456;

Best Local Similarity 100.0%; Pred. No. 2.6e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAATAG 60
 |||
 Db 173 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAATAG 114
 |||
 Qy 61 GGGTTCGCGCACATTTCCCGGAAAAGTGCCACCTGACGTC 101
 |||
 Db 113 GGGTTCGCGCACATTTCCCGGAAAAGTGCCACCTGACGTC 73
 |||

RESULT 6

AAN71080/c
 ID AAN71080 standard; DNA; 456 BP.
 XX
 AC AAN71080;
 XX
 DT 25-MAR-2003 (revised)
 DT 10-MAR-2003 (revised)
 DT 13-MAY-1991 (first entry)
 XX
 DE Sequence encoding beta-urogastrone.
 XX
 KW pUGT 150s; beta-UG; ds.
 XX
 OS Escherichia coli.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT promoter 125..170
 FT /*tag= a
 FT CDS 209..439
 FT /*tag= b
 FT /transl_except= (pos:434..436,aa:Arg)
 FT
 XX JPG2190083-A.
 XX
 XX 20-AUG-1987.
 XX
 PF 14-FEB-1986; 86JP-00031415.
 XX
 PR 14-FEB-1986; 86JP-00031415.
 XX
 XX (EART) EARTH SEIYAKU KK.
 XX
 XX WPI; 1987-273761/39.
 DR
 XX Expression vector contg. multiple information units is used to transform host all for increased prodn. of polypeptides.
 FT
 XX Disclosure; Page 553; 34pp; Japanese.
 XX
 CC Sequence encodes beta-urogastrone under the control of a tac promoter. The peptide may be expressed from plasmid pUGT 150s in a transformed E.coli host. The plasmid may carry several separately expressing sequences comprising a tac promoter, SD site, signal peptide, and coding sequence, to produce beta-UG in high yield. (Updated on 10-MAR-2003 to add missing OS field.) (Updated on 25-MAR-2003 to correct PA field.)
 XX
 SQ Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;
 Query Match 100.0%; Score 101; DB 1; Length 456;
 Best Local Similarity 100.0%; Pred. No. 2.6e-21;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAATAG 60
 |||
 Db 173 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAATAG 114
 |||
 Qy 61 GGGTTCGCGCACATTTCCCGGAAAAGTGCCACCTGACGTC 101
 |||
 Db 113 GGGTTCGCGCACATTTCCCGGAAAAGTGCCACCTGACGTC 73
 |||

RESULT 7

AAN70833/c
 ID AAN70833 standard; DNA; 456 BP.
 XX
 AC AAN70833;
 XX
 DT 25-MAR-2003 (revised)
 DT 10-MAR-2003 (revised)
 DT 18-JAN-1991 (first entry)
 XX
 DE Beta-urogastrone sequence.
 XX
 KW Tumour; inosine; DNA probe; ds.
 XX
 OS Unidentified.
 XX
 FH Key Location/Qualifiers
 FT promoter 125..170
 FT /*tag= b
 FT RBS 200..204
 FT /*tag= c
 FT CDS 209..439
 FT /*tag= a
 FT sig_peptide 209..277
 FT /*tag= d
 XX
 XX JP62244398-A.
 XX
 XX 24-OCT-1987.
 XX
 PF 16-APR-1986; 86JP-00087368.
 XX
 PR 16-APR-1986; 86JP-00087368.
 XX
 XX (SEKI) SEKISUI CHEM IND CO LTD.
 XX
 XX WPI; 1987-339045/48.
 DR
 DR P-PSDB; AAP70505.
 XX
 XX Detection of DNA and/or RNA - by converting to single strand form and using probe contg. labelled inosine deriv.
 FT
 XX Disclosure; Page 11; 11pp; Japanese.
 XX
 CC An example of a sequence detected by a probe consisting of polyinosine, polydeoxyinosine, oligoinosine and/or oligodeoxyinosine is labeled. The ssDNA and probe are hybridized and the existence of DNA in the product is detected. It can be used to detect the presence of malignant tumour.
 CC (Updated on 10-MAR-2003 to add missing OS field.) (Updated on 25-MAR-2003 to correct PA field.)
 XX
 SQ Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;
 Query Match 100.0%; Score 101; DB 1; Length 456;
 Best Local Similarity 100.0%; Pred. No. 2.6e-21;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAATAG 60
 |||
 Db 173 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAATAG 114
 |||
 Qy 61 GGGTTCGCGCACATTTCCCGGAAAAGTGCCACCTGACGTC 101
 |||
 Db 113 GGGTTCGCGCACATTTCCCGGAAAAGTGCCACCTGACGTC 73
 |||
 RESULT 8
 AAN81765/c
 ID AAN81765 standard; DNA; 456 BP.
 XX
 AC AAN81765;
 XX
 DT 25-MAR-2003 (revised)
 DT 13-DEC-1990 (first entry)


```
XX Sequence of HB101 (pGH201) encoding new beta-urogastrone deriv. Met (21),
DE Arg (53).
XX
XX Gastric acid secretion; cell proliferation; hormone; ds.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX CDS 209..277
XX FT /*tag= a
XX FT 278..439
XX FT /*tag= b
XX FT /product= "New beta-urogastrone deriv."
XX
XX JP63012298-A.
XX
XX 19-JAN-1988.
XX
XX 30-JUN-1986; 86JP-00153783.
XX
XX 30-JUN-1986; 86JP-00153783.
XX
XX (EART ) EARTH SEIYAKU KK.
XX
XX WPI; 1988-054638/08.
XX DR P-PSDB; AAP81349.
XX
XX New beta-urogastrone deriv. - has gastric acid secretion inhibition and
XX proliferation promotion activity.
XX
XX Disclosure; Page 685; 76pp; Japanese.
XX
XX The deriv. has various biological activities such as gastric acid
XX secretion inhibiting action, or cell proliferation promoting action. The
XX deriv. has the same biological or pharmacological activities as beta-
XX urogastrone. It is not susceptible to denaturation by oxidn. and is
XX chemically stable. Deriv. has resistance to proteolytic enzymes such as
XX protease. (Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 101; DB 1; Length 456;
XX Best Local Similarity 100.0%; Pred. No. 2.6e-21;
XX Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTAGAAAATAAACAAATAG 60
XX 173 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTAGAAAATAAACAAATAG 114
XX
XX 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
XX 113 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 73
XX
XX RESULT 9
XX ABA90413/c
XX ID ABA90413 standard; DNA; 466 BP.
XX
XX AC ABA90413;
XX
XX 12-FEB-2002 (first entry)
XX
XX Drosophila cell cycle progression protein coding sequence #48.
XX
XX Antiproliferative; cytostatic; cardiant; immunosuppressive; meiosis;
XX antiinflammatory; antiparasitic; dermatologic; antifungal; mitosis;
XX antiparasitic; antimalarial; antirheumatic; antiarthritic; cell division;
XX cell cycle progression protein; tumour; proliferative disorder;
XX cardiovascular; autoimmune; dermatological disorder; ds.
XX
XX Drosophila sp.
XX
XX Sequence 466 BP; 135 A; 87 C; 93 G; 150 T; 0 U; 1 Other;
XX
XX Query Match 100.0%; Score 101; DB 6; Length 466;
XX Best Local Similarity 100.0%; Pred. No. 2.6e-21;
XX Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTAGAAAATAAACAAATAG 60
XX 280 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTAGAAAATAAACAAATAG 221
XX
XX 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
XX 220 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 180
XX
XX RESULT 10
XX AAX211173/c
XX ID AAX211173 standard; DNA; 487 BP.
XX
XX AC AAX211173;
XX
XX 05-MAY-1999 (first entry)
XX
XX Polynucleotide sequence from the genome of Treponema pallidum.
XX
XX Treponema pallidum infection; syphilis; Borrelia infection; animal;
XX enzyme production; ds.
XX
XX Treponema pallidum.
XX
XX OS WO9859034-A2.
XX
XX 30-DEC-1998.
XX
XX 23-JUN-1998; 98WO-US013041.
XX
XX 24-JUN-1997; 97US-0050667P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Fraser CM;
XX
```

DR WPI; 1999-081273/07.

XX New isolated *Treponema pallidum* nucleic acids - used to develop products

PT for the detection, diagnosis, characterisation, prevention and therapy of

PT *T. pallidum* infections, particularly syphilis.

XX

PS Claim 1; Page 1106; 1150pp; English.

XX

CC AAX20500-21243 represent polynucleotide sequences from the genome of

CC *Treponema pallidum*. The sequences can be used for detection, diagnosis,

CC characterisation, prevention and therapy for *T. pallidum* infections,

CC particularly syphilis. They can also be used for detecting diseases

CC related to *Borrelia* infections in animals, and for the production of

CC biosynthetic products such as enzymes

XX

SQ Sequence 487 BP; 125 A; 127 C; 113 G; 121 T; 0 U; 1 Other;

Query Match 100.0%; Score 101; DB 2; Length 487;

Best Local Similarity 100.0%; Pred. No. 2.6e-21;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 60

DB 323 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 264

QY 61 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGACGTC 101

DB 263 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGACGTC 223

RESULT 11

AAX21149/C

ID AAX21149 standard; DNA; 535 BP.

XX

AC AAX21149;

XX

XX 05-MAY-1999 (first entry)

XX

DE Polynucleotide sequence from the genome of *Treponema pallidum*.

XX

XX *Treponema pallidum* infection; syphilis; *Borrelia* infection; animal;

KW enzyme production; ds.

XX

XX *Treponema pallidum*.

OS

XX WO9859034-A2.

XX

XX 30-DEC-1998.

XX

XX 23-JUN-1998; 98WO-US013041.

XX

XX 24-JUN-1997; 97US-0050667P.

XX

PA (HUMA-) HUMAN GENOME SCI INC.

XX

PI Fraser CM;

XX

DR WPI; 1999-081273/07.

XX

XX New isolated *Treponema pallidum* nucleic acids - used to develop products

PT for the detection, diagnosis, characterisation, prevention and therapy of

PT *T. pallidum* infections, particularly syphilis.

XX

PS Claim 1; Page 1093; 1150pp; English.

XX

CC AAX20500-21243 represent polynucleotide sequences from the genome of

CC *Treponema pallidum*. The sequences can be used for detection, diagnosis,

CC characterisation, prevention and therapy for *T. pallidum* infections,

CC particularly syphilis. They can also be used for detecting diseases

CC related to *Borrelia* infections in animals, and for the production of

CC biosynthetic products such as enzymes

XX

SQ Sequence 535 BP; 145 A; 108 C; 122 G; 155 T; 0 U; 5 Other;

Query Match 100.0%; Score 101; DB 2; Length 535;

Best Local Similarity 100.0%; Pred. No. 2.7e-21;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 60

DB 355 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 296

RESULT 12

ABA90456/c

ID ABA90456 standard; DNA; 573 BP.

XX

AC ABA90456;

XX

XX 12-FEB-2002 (first entry)

XX

DE *Drosophila* cell cycle progression protein coding sequence #91.

XX

XX Antiproliferative; cytostatic; cardiant; immunosuppressive; meiosis;

KW antiinflammatory; antipsoriatic; dermatological; antifungal; mitosis;

KW antiparasitic; antimalarial; antirheumatic; antiarthritic; cell division;

KW cell cycle progression protein; tumour; proliferative disorder;

KW cardiovascular; autoimmune; dermatological disorder; ds.

XX

OS *Drosophila* sp.

XX

XX WO200172774-A2.

XX

XX 04-OCT-2001.

XX

XX 23-MAR-2001; 2001WO-GB001297.

XX

XX 24-MAR-2000; 2000GB-00007268.

XX

PA (CYCL-) CYCLACEL LTD.

XX

XX Deak P, Glover DM, Midgley C;

XX

XX WPI; 2002-055132/07.

XX

XX Polynucleotides encoding cell cycle progression proteins, useful for

PT treating a tumor or a proliferative disorder.

XX

PS Claim 1; Page 144; 213pp; English.

XX

CC The present invention relates to *Drosophila* cell cycle progression

CC proteins (AAM47572-AAM47608) and their coding sequences (ABA90366-

CC ABA90520). The coding sequences and proteins are useful for identifying a

CC substance capable of affecting the function of the corresponding gene, a

CC substance capable of inhibiting the cell division cycle, or capable of

CC inhibiting mitosis and/or meiosis. They can also be used in a method for

CC treating a tumor or proliferative disorder, cardiovascular disorders

CC (such as restenosis and cardiomyopathy), autoimmune disorders such as

CC (glomerulonephritis and rheumatoid arthritis), dermatological disorders

CC (such as psoriasis), antiinflammatory, antifungal and antiparasitic

CC disorders (such as malaria)

XX

SQ Sequence 573 BP; 154 A; 118 C; 116 G; 184 T; 0 U; 1 Other;

Query Match 100.0%; Score 101; DB 6; Length 573;

Best Local Similarity 100.0%; Pred. No. 2.7e-21;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 60

DB 355 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 296

PR	06-SEP-2000;	2000US-0230438P
PR	08-SEP-2000;	2000US-0231242P
PR	08-SEP-2000;	2000US-0231243P
PR	08-SEP-2000;	2000US-0231243P
PR	08-SEP-2000;	2000US-0231244P
PR	08-SEP-2000;	2000US-0231413P
PR	08-SEP-2000;	2000US-0231414P
PR	08-SEP-2000;	2000US-0232080P
PR	08-SEP-2000;	2000US-0232081P
PR	12-SEP-2000;	2000US-0231968P
PR	14-SEP-2000;	2000US-0232397P
PR	14-SEP-2000;	2000US-0232398P
PR	14-SEP-2000;	2000US-0232399P
PR	14-SEP-2000;	2000US-0232400P
PR	14-SEP-2000;	2000US-0234011P
PR	14-SEP-2000;	2000US-0233063P
PR	14-SEP-2000;	2000US-0233064P
PR	14-SEP-2000;	2000US-0233065P
PR	21-SEP-2000;	2000US-0234223P
PR	21-SEP-2000;	2000US-0234224P
PR	21-SEP-2000;	2000US-0234225P
PR	25-SEP-2000;	2000US-0234997P
PR	25-SEP-2000;	2000US-0234998P
PR	26-SEP-2000;	2000US-0234548P
PR	27-SEP-2000;	2000US-0233834P
PR	27-SEP-2000;	2000US-0233836P
PR	29-SEP-2000;	2000US-0236327P
PR	29-SEP-2000;	2000US-0236367P
PR	29-SEP-2000;	2000US-0236368P
PR	29-SEP-2000;	2000US-0236369P
PR	29-SEP-2000;	2000US-0236370P
PR	02-OCT-2000;	2000US-0237038P
PR	02-OCT-2000;	2000US-0237039P
PR	02-OCT-2000;	2000US-0237040P
PR	13-OCT-2000;	2000US-0239935P
PR	13-OCT-2000;	2000US-0239937P
PR	20-OCT-2000;	2000US-0240960P
PR	20-OCT-2000;	2000US-0241211P
PR	20-OCT-2000;	2000US-0241809P
PR	20-OCT-2000;	2000US-0241826P
PR	01-NOV-2000;	2000US-0246171P
PR	01-NOV-2000;	2000US-0246474P
PR	08-NOV-2000;	2000US-0246475P
PR	08-NOV-2000;	2000US-0246476P
PR	08-NOV-2000;	2000US-0246477P
PR	08-NOV-2000;	2000US-0246478P
PR	08-NOV-2000;	2000US-0246523P
PR	08-NOV-2000;	2000US-0246524P
PR	08-NOV-2000;	2000US-0246525P
PR	08-NOV-2000;	2000US-0246526P
PR	08-NOV-2000;	2000US-0246527P
PR	08-NOV-2000;	2000US-0246528P
PR	08-NOV-2000;	2000US-0246532P
PR	08-NOV-2000;	2000US-0246609P
PR	08-NOV-2000;	2000US-0246610P
PR	08-NOV-2000;	2000US-0246611P
PR	08-NOV-2000;	2000US-0246613P
PR	17-NOV-2000;	2000US-0249207P
PR	17-NOV-2000;	2000US-0249208P
PR	17-NOV-2000;	2000US-0249209P
PR	17-NOV-2000;	2000US-0249210P
PR	17-NOV-2000;	2000US-0249211P
PR	17-NOV-2000;	2000US-0249212P
PR	17-NOV-2000;	2000US-0249213P
PR	17-NOV-2000;	2000US-0249214P
PR	17-NOV-2000;	2000US-0249215P
PR	17-NOV-2000;	2000US-0249216P
PR	17-NOV-2000;	2000US-0249217P
PR	17-NOV-2000;	2000US-0249218P

PR	17-NOV-2000;	2000US-0249244P.
PR	17-NOV-2000;	2000US-0249244P.
PR	17-NOV-2000;	2000US-0249245P.
PR	17-NOV-2000;	2000US-0249246P.
PR	17-NOV-2000;	2000US-0249265P.
PR	17-NOV-2000;	2000US-0249265P.
PR	17-NOV-2000;	2000US-0249277P.
PR	17-NOV-2000;	2000US-0249299P.
PR	17-NOV-2000;	2000US-0249300P.
PR	01-DEC-2000;	2000US-0250160P.
PR	01-DEC-2000;	2000US-0250319P.
PR	05-DEC-2000;	2000US-0251030P.
PR	05-DEC-2000;	2000US-0251198P.
PR	05-DEC-2000;	2000US-0256719P.
PR	06-DEC-2000;	2000US-0251479P.
PR	08-DEC-2000;	2000US-0251856P.
PR	08-DEC-2000;	2000US-0251868P.
PR	08-DEC-2000;	2000US-0251869P.
PR	08-DEC-2000;	2000US-0251989P.
PR	08-DEC-2000;	2000US-0251990P.
PR	11-DEC-2000;	2000US-0254097P.
PR	05-JAN-2001;	2001US-0259678P.
XX	(HUMA-) HUMAN	GENOME SCI INC.
PA		

Novel isolated prostate gland related polypeptide useful for diagnosis and treatment of disorders of prostate such as prostatodystonia, prostatosis, prostatitis, benign prostatic hypertrophy and malacoplakia

Claim 1: SEQ ID NO 418: 512bp: English.

The invention relates to novel isolated prostate gland related nucleic acids (I) and polypeptides (II). (I) and (II) are useful for diagnosis, prognosis, prevention, and/or treatment of diseases and/or disorders of the prostate such as acute non-bacterial prostatitis, chronic non-bacterial prostatitis, acute bacterial prostatitis, prostatic dysplasia, prostatic adenocarcinomas, transitional cell carcinomas, ductal carcinomas, and squamous cell carcinomas. (I), (II) and antibody to (II) are useful for diagnosing and treating reproductive system disorders (Paget's disease), autoimmune disorders (systemic lupus erythematosus, rheumatoid arthritis), blood-related disorders (sickle cell anaemia), hyperproliferative disorders, urinary system disorders (glomerulonephritis), cardiovascular disorders (arrhythmias), respiratory disorders, musculoskeletal system disorders, neural activity and neurological disorders (Alzheimer's disease and Parkinson's disease), endocrine disorders (Addison's disease), gastrointestinal disorders (inflammatory disorders), liver disorders (biliary liver cirrhosis), pancreatic and gall bladder disorders, disorders of the large intestine, developmental and inherited disorders, diseases at the cellular level, and wound healing and epithelial cell proliferation. (I) or (II) is useful to prevent skin aging, for preventing hair loss, to maintain organs before transplantation, and as food additive or preservative.

Query Match	Score 101;	DB 4;	Length 776;
Best Local Similarity	100.0%;	Pred. No. 2.9e-21;	
Matches 101;	Conservative	0;	Mismatches 0;
	Indels	0;	Gaps 0;

QY	1	AGGGTTATTGTCTCATGAGCGGATACATATTTTGAATGTATTTAGAAAAATAAACCAATAG	60
Db	546	AGGGTTATTGTCTCATGAGCGGATACATATTTTGAATGTATTTAGAAAAATAAACCAATAG	487
QY	61	GGGTTCCGCGACATTTCCCGAAAAAGTGCCACCTGACGTC	101
Dp	486	GGGTTCCGCGACATTTCCCGAAAAAGTGCCACCTGACGTC	446

RESULT 15
AAS27819/C
ID AAS27819 standard; DNA; 776 BP.

PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Barash SC, Ruben SM;
 PI
 XX WPI; 2001-465460/50.
 XX
 XX Novel polypeptides useful for diagnosing, treating, preventing and/or
 PT prognosing disorders related to the proteins, including cancers, immune
 PT disorders and neuronal disorders.
 XX
 XX Claim 1; SEQ ID NO 1479; 880pp; English.
 XX
 XX The invention relates to novel isolated polypeptides (I), and
 CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for
 CC diagnosing, preventing and treating diseases including immune system
 CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
 CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
 CC transplant rejections and graft versus host disease, infectious diseases
 CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
 CC other blood-related disorders (sickle cell anaemia), myeloproliferative
 CC disorders, primary haematopoietic disorders, hyperproliferative disorders
 CC (e.g. Gaucher's disease and cancer), neurodegenerative disorders (e.g.
 CC Alzheimer's disease, Parkinson's disease), chromosomal abnormalities
 CC (Down syndrome), ischaemic injury (e.g. stroke), renal disorders (e.g.
 CC glomerulonephritis), cardiovascular disorders (e.g. arrhythmia),
 CC respiratory disorders, dermatological disorders, in wound healing,
 CC epithelial cell proliferation, endocrine disorders (e.g. Addison's
 CC disease), reproductive system disorders, gastrointestinal disorder
 CC (inflammatory disorders), liver disorders (cirrhosis), as stimulators of
 CC B-cell responsiveness to pathogens, activators of T-cells, to induce
 CC higher affinity antibodies, and as a means to induce tumour proliferation
 CC in pathologies e.g. acquired immune deficiency syndrome (AIDS). AAS26976-
 CC AAS27850 represent novel signal transduction pathway protein coding
 CC sequences and PCR primers of the invention
 XX
 XX Sequence 776 BP; 184 A; 209 C; 181 G; 200 T; 0 U; 2 Other;
 SQ
 Query Match 100.0%; Score 101; DB 4; Length 776;
 Best Local Similarity 100.0%; Pred. No. 2.9e-21;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AGGGTTATTCCTCATGACGGGATACATATTTTGAATGATTTAGAAAAATAACAAATAG 60
 DB 546 AGGGTTATTCCTCATGACGGGATACATATTTTGAATGATTTAGAAAAATAACAAATAG 487
 QY 61 GGGGTTCGGGCACATTTCCCGAAAAATGCGCACTGACGTC 101
 Db 486 GGGGTTCGGGCACATTTCCCGAAAAATGCGCACTGACGTC 446

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 05:15:57 ; Search time 961.667 Seconds
(without alignments)
3997.736 Million cell updates/sec

Title: US-09-482-682-44_COPY_7860_7960
Perfect score: 101
Sequence: 1 agggttattgtctcatgagc.....gaaagtgccactgcagtc 101

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*

1: gb_est1.*
2: gb_est2.*
3: gb_hic.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gss1.*
9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	101	100.0	300	4	BM078095
C 2	101	100.0	300	5	BU963956
C 3	101	100.0	300	5	BU964094
C 4	101	100.0	309	9	AL000426
C 5	101	100.0	391	1	AL597149
C 6	101	100.0	414	9	CC819240
C 7	101	100.0	417	4	BJ684174
C 8	101	100.0	491	9	CC819923
C 9	101	100.0	495	4	B1805285
C 10	101	100.0	495	9	CC818374
C 11	101	100.0	496	9	CC818523
C 12	101	100.0	503	9	CC819854
C 13	101	100.0	515	9	CC817752
C 14	101	100.0	518	9	CC817128
C 15	101	100.0	519	9	CC817162
C 16	101	100.0	519	9	CC817796
C 17	101	100.0	521	9	CC819067
C 18	101	100.0	533	9	CC819841
C 19	101	100.0	542	9	CC816892
C 20	101	100.0	550	7	CR766622
C 21	101	100.0	551	9	CC816905
C 22	101	100.0	554	9	CC819058
C 23	101	100.0	563	9	CC819270
C 24	101	100.0	566	9	CC816848

C 25	101	100.0	566	9	CC820024
C 26	101	100.0	567	9	CC817070
C 27	101	100.0	568	9	CC818640
C 28	101	100.0	571	9	CC816954
C 29	101	100.0	571	9	CC818423
C 30	101	100.0	580	9	CC819098
C 31	101	100.0	582	1	AL694813
C 32	101	100.0	583	9	CC817633
C 33	101	100.0	583	9	CC818436
C 34	101	100.0	586	9	CC816883
C 35	101	100.0	588	9	CC817788
C 36	101	100.0	588	9	CC818340
C 37	101	100.0	589	9	CC817595
C 38	101	100.0	590	9	CC819754
C 39	101	100.0	592	9	CC817679
C 40	101	100.0	592	9	CC818508
C 41	101	100.0	593	9	CC816942
C 42	101	100.0	593	9	CC817699
C 43	101	100.0	594	9	CC818287
C 44	101	100.0	594	9	CC818422
C 45	101	100.0	595	9	CC816929

ALIGNMENTS

RESULT 1
BM078095/c
LOCUS 300 bp mRNA linear EST 30-NOV-2001
DEFINITION 83374 Hebeloma cylindrosporum functional cDNA library Hebeloma cylindrosporum cDNA 5', mRNA sequence.
ACCESSION BM078095
VERSION BM078095.1 GI:17157967
KEYWORDS EST.
SOURCE Hebeloma cylindrosporum
ORGANISM Hebeloma cylindrosporum
Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes; Agaricales; Cortinariaceae; Hebeloma.
REFERENCE 1 (bases 1 to 300)
AUTHORS Wipf D., Benjdia, M., Tegeder, M. and Frommer, W.B.
TITLE Construction of a functional cDNA library from the ectomycorrhizal fungus Hebeloma cylindrosporum
JOURNAL Unpublished (2001)
COMMENT Contact: Wipf D
ZMBP - Center for Molecular Biology of Plants
University of Tuebingen
Auf der Morgenstelle 1, 72076 Tuebingen, Germany
Tel: 49 7071 2976160
Fax: 49 7071 293287
Email: daniel.wipf@zmbp.uni-tuebingen.de
PCR Primers
FORWARD: pDR196 5' primer (PMA 5')
High quality sequence stop: 300
POLYA=No.

FEATURES
source
1..300
Location/Qualifiers
/organism="Hebeloma cylindrosporum"
/mol_type="mRNA"
/strain="H1"
/db_xref="taxon:76867"
/tissue_type="Mycelia"
/lab_host="E. coli XLI-Blue"
/clone_lib="Hebeloma cylindrosporum functional cDNA library"
/note="vector: pDR 196 (unpublished); Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Query Match 100.0%; Score 101; DB 4; Length 300;
Best Local Similarity 100.0%; Pred. No. 8.1e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGTTATTGCTCATGCGGATACATATTGTAATGCTATTAGAAAAATAACAATAG 60

```

|||||
174 AGGGTTATTGCTCATGCGGACATATTTGAATGTTATTAGAAAAATAACAATAG 115
|||||
Qy 61 GGGTTCCGCGCACATTTCCCGGAAAGTGCACCTGACGTC 101
|||||
Db 114 GGGTTCCGCGCACATTTCCCGGAAAGTGCACCTGACGTC 74
|||||

RESULT 2
BU963956/c
LOCUS EST88 Hebeloma cylindrosporum functional cDNA library Hebeloma
DEFINITION cylindrosporum cDNA, mRNA sequence.
ACCESSION BU963956
VERSION BU963956.1 GI:24204753
KEYWORDS EST.
SOURCE Hebeloma cylindrosporum
ORGANISM Hebeloma cylindrosporum
Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
Agaricales; Cortinariaceae; Hebeloma.
REFERENCE 1 (bases 1 to 300)
AUTHORS Wipf, D., Benjdia, M., Rikirsch, E., Zimmermann, S., Tegeder, M. and
Frommer, W.B.
TITLE A Hebeloma cylindrosporum expression cDNA library for suppression
cloning in yeast mutants, complementation of a yeast his4 mutant
and EST analysis
JOURNAL Unpublished (2002)
COMMENT Contact: Wipf D
ZMBP - Center for Molecular Biology of Plants
University of Tuebingen
Auf der Morgenstelle 1, 72076 Tuebingen, Germany
Tel: 49 7071 2976160
Fax: 49 7071 293287
Email: daniel.wipf@zmbp.uni-tuebingen.de
homolog to Enterobacteria phage f1 ; ampicillinase (1e-10) .
FEATURES
source
1..300
/organism="Hebeloma cylindrosporum"
/mol_type="mRNA"
/strain="H1"
/db_xref="taxon:76867"
/tissue_type="Mycelia"
/lab_host="E. coli XL1-Blue"
/clone_lib="Hebeloma cylindrosporum functional cDNA
library"
/notes="Vector: pDR 196 (unpublished); Site_1: EcoRI;
Site_2: XhoI"

ORIGIN
Query Match 100.0%; Score 101; DB 5; Length 300;
Best Local Similarity 100.0%; Pred. No. 8.1e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGCGGACATATTTGAATGTTATTAGAAAAATAACAATAG 60
|||||
Db 170 AGGGTTATTGCTCATGCGGACATATTTGAATGTTATTAGAAAAATAACAATAG 111
|||||

Qy 61 GGGTTCCGCGCACATTTCCCGGAAAGTGCACCTGACGTC 101
|||||
Db 110 GGGTTCCGCGCACATTTCCCGGAAAGTGCACCTGACGTC 70
|||||

RESULT 4
FR0009140
LOCUS F.rubripes GSS sequence, clone 010H20ac4, genomic survey sequence.
DEFINITION F.rubripes GSS sequence, clone 010H20ac4, genomic survey sequence.
ACCESSION AL000426
VERSION AL000426.1 GI:2438278
KEYWORDS GSS; genome survey sequence.
SOURCE Takifugu rubripes (Fugu rubripes)
ORGANISM Takifugu rubripes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Takifugu.
1
REFERENCE 1
AUTHORS Elgar, G., Clark, M.S., Meek, S., Smith, S., Warner, S., Edwards, Y.J.,
Bouchireb, N., Cottage, A., Yeo, G.S., Umrانيا, Y., Williams, G. and
Brenner, S.
TITLE Generation and analysis of 25 Mb of genomic DNA from the pufferfish
Fugu rubripes by sequence scanning
JOURNAL Genome Res. 9 (10), 960-971 (1999)
MEDLINE 99455097
PUBMED 10523524
REFERENCE 2 (bases 1 to 309)
AUTHORS Elgar, G., Clark, M., Smith, S., Meek, S., Warner, S., Umrانيا, Y.,
Williams, G. and Brenner, S.
TITLE Direct Submission
JOURNAL Submitted (09-SEP-1997) MRC Human Genome Mapping Project Resource
Centre Hinxton, Cambridge, CB10 1SB. Email: biohelp@hgm.mrc.ac.uk
COMMENT Vector: pBluescript II KS
V_type: phagemid

```

```

|||||
174 AGGGTTATTGCTCATGCGGACATATTTGAATGTTATTAGAAAAATAACAATAG 115
|||||
Qy 61 GGGTTCCGCGCACATTTCCCGGAAAGTGCACCTGACGTC 101
|||||
Db 114 GGGTTCCGCGCACATTTCCCGGAAAGTGCACCTGACGTC 74
|||||

RESULT 2
BU963956
LOCUS EST88 Hebeloma cylindrosporum functional cDNA library Hebeloma
DEFINITION cylindrosporum cDNA, mRNA sequence.
ACCESSION BU963956
VERSION BU963956.1 GI:24204753
KEYWORDS EST.
SOURCE Hebeloma cylindrosporum
ORGANISM Hebeloma cylindrosporum
Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
Agaricales; Cortinariaceae; Hebeloma.
REFERENCE 1 (bases 1 to 300)
AUTHORS Wipf, D., Benjdia, M., Rikirsch, E., Zimmermann, S., Tegeder, M. and
Frommer, W.B.
TITLE A Hebeloma cylindrosporum expression cDNA library for suppression
cloning in yeast mutants, complementation of a yeast his4 mutant
and EST analysis
JOURNAL Unpublished (2002)
COMMENT Contact: Wipf D
ZMBP - Center for Molecular Biology of Plants
University of Tuebingen
Auf der Morgenstelle 1, 72076 Tuebingen, Germany
Tel: 49 7071 2976160
Fax: 49 7071 293287
Email: daniel.wipf@zmbp.uni-tuebingen.de
no homology below 1e-10.
FEATURES
source
1..300
/organism="Hebeloma cylindrosporum"
/mol_type="mRNA"
/strain="H1"
/db_xref="taxon:76867"
/tissue_type="Mycelia"
/lab_host="E. coli XL1-Blue"
/clone_lib="Hebeloma cylindrosporum functional cDNA
library"
/notes="Vector: pDR 196 (unpublished); Site_1: EcoRI;
Site_2: XhoI"

ORIGIN
Query Match 100.0%; Score 101; DB 5; Length 300;
Best Local Similarity 100.0%; Pred. No. 8.1e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGCGGACATATTTGAATGTTATTAGAAAAATAACAATAG 60
|||||
Db 170 AGGGTTATTGCTCATGCGGACATATTTGAATGTTATTAGAAAAATAACAATAG 111
|||||

Qy 61 GGGTTCCGCGCACATTTCCCGGAAAGTGCACCTGACGTC 101
|||||
Db 110 GGGTTCCGCGCACATTTCCCGGAAAGTGCACCTGACGTC 70
|||||

RESULT 4
FR0009140
LOCUS F.rubripes GSS sequence, clone 010H20ac4, genomic survey sequence.
DEFINITION F.rubripes GSS sequence, clone 010H20ac4, genomic survey sequence.
ACCESSION AL000426
VERSION AL000426.1 GI:2438278
KEYWORDS GSS; genome survey sequence.
SOURCE Takifugu rubripes (Fugu rubripes)
ORGANISM Takifugu rubripes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Takifugu.
1
REFERENCE 1
AUTHORS Elgar, G., Clark, M.S., Meek, S., Smith, S., Warner, S., Edwards, Y.J.,
Bouchireb, N., Cottage, A., Yeo, G.S., Umrانيا, Y., Williams, G. and
Brenner, S.
TITLE Generation and analysis of 25 Mb of genomic DNA from the pufferfish
Fugu rubripes by sequence scanning
JOURNAL Genome Res. 9 (10), 960-971 (1999)
MEDLINE 99455097
PUBMED 10523524
REFERENCE 2 (bases 1 to 309)
AUTHORS Elgar, G., Clark, M., Smith, S., Meek, S., Warner, S., Umrانيا, Y.,
Williams, G. and Brenner, S.
TITLE Direct Submission
JOURNAL Submitted (09-SEP-1997) MRC Human Genome Mapping Project Resource
Centre Hinxton, Cambridge, CB10 1SB. Email: biohelp@hgm.mrc.ac.uk
COMMENT Vector: pBluescript II KS
V_type: phagemid

```



```

PRIMER: KS
DESCR:
One pass dye-terminator sequencing of cosmid cloned genomic
sequence.
FEATURES
    source
        Location/Qualifiers
            1..309
                /organism="Takifugu rubripes"
                /mol_type="genomic DNA"
                /db_xref="taxon:31033"
                /clone="010H20aC4"
                /clone_lib="cosmid 010H20"
ORIGIN
Query Match      100.0%; Score 101; DB 9; Length 309;
Best Local Similarity 100.0%; Pred. No. 8.2e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGTTATTGTCATGAGCGGATACATATTGAATGTTATTAGAAAATAAACAATAG 60
Db 39 AGGTTATTGTCATGAGCGGATACATATTGAATGTTATTAGAAAATAAACAATAG 98
Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
Db 99 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 139
RESULT 5
AL597149
LOCUS
DEFINITION DKFZp313J1611_r1 313 (synonym: hlcc2) Homo sapiens cDNA clone
ACCESSION AL597149
VERSION AL597149.1 GI:15154845
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Koehrer,K., Beyer,A., Mewes,W., Weil,B. and Wiemann,S.
TITLE EST (Koehrer,K., Beyer,A., Mewes,H.W., Weil,B. and Wiemann,S.)
JOURNAL Unpublished (1999)
COMMENT Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert.
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by BMFZ (Biomedical Research Center at the Charite,
Berlin/Germany) within the cDNA sequencing consortium of the German
Genome Project.
No sl sequence available.
This clone (DKFZp313J1611) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
    source
        Location/Qualifiers
            1..391
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="DKFZp313J1611"
                /dev_stage="adult"
                /lab_host="DH10B"
                /clone_lib="313 (synonym: hlcc2)"
                /note="Vector: pTriplex2; Site_1: sf1A; Site_2: sf1B;
                cDNA-collection"
ORIGIN
Query Match      100.0%; Score 101; DB 1; Length 391;
Best Local Similarity 100.0%; Pred. No. 8.3e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGTTATTGTCATGAGCGGATACATATTGAATGTTATTAGAAAATAAACAATAG 60
Db 414 AGGTTATTGTCATGAGCGGATACATATTGAATGTTATTAGAAAATAAACAATAG 355
PRIMER: KS
DESCR:
One pass dye-terminator sequencing of cosmid cloned genomic
sequence.
FEATURES
    source
        Location/Qualifiers
            1..309
                /organism="Takifugu rubripes"
                /mol_type="genomic DNA"
                /db_xref="taxon:31033"
                /clone="010H20aC4"
                /clone_lib="cosmid 010H20"
ORIGIN
Query Match      100.0%; Score 101; DB 9; Length 309;
Best Local Similarity 100.0%; Pred. No. 8.2e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGTTATTGTCATGAGCGGATACATATTGAATGTTATTAGAAAATAAACAATAG 60
Db 39 AGGTTATTGTCATGAGCGGATACATATTGAATGTTATTAGAAAATAAACAATAG 98
Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
Db 99 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 139
RESULT 6
CC819240
LOCUS
DEFINITION 10005D19R Oxytricha plasmid UUC10 library Sterkiella
histriomuscorum genomic clone UUC10005D19 R, genomic survey
sequence.
ACCESSION CC819240
VERSION CC819240.1 GI:32899308
KEYWORDS GSS.
SOURCE Sterkiella histriomuscorum (Oxytricha trifallax)
ORGANISM Sterkiella histriomuscorum
REFERENCE Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
Stichotrichida; Oxytrichidae; Sterkiella.
AUTHORS Dunn,D., Doak,T., Herrick,G. and Weiss,R.
TITLE Paired end reads from plasmid inserts of Oxytricha trifallax
macronuclear chromosomes
JOURNAL Unpublished (2003)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Plate: 0005 row: D column: 19
Seq primer: CACACAGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 414.
FEATURES
    source
        Location/Qualifiers
            1..414
                /organism="Sterkiella histriomuscorum"
                /mol_type="genomic DNA"
                /db_xref="taxon:94289"
                /clone="UUC10005D19"
                /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                /clone_lib="Oxytricha plasmid UUC10 library"
                /note="Vector: PWD42nv; Purified macronuclear chromosomal
                DNA from Oxytricha trifallax was blunt end-repaired with
                T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
                oligonucleotides were ligated to the blunt ends in high
                molar excess. Vector DNA was prepared from a derivative of
                PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
                derivative of plasmid R1. The vector was ligated with
                adaptors complementary to the insert adaptors and
                purified. The sheared, adapted mouse DNA was annealed to
                adapted vector DNA, and transformed into
                chemically-competent E. Coli XL10-Gold (Stratagene) cells
                and selected for ampicillin resistance."
ORIGIN
Query Match      100.0%; Score 101; DB 9; Length 414;
Best Local Similarity 100.0%; Pred. No. 8.3e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGTTATTGTCATGAGCGGATACATATTGAATGTTATTAGAAAATAAACAATAG 60
Db 414 AGGTTATTGTCATGAGCGGATACATATTGAATGTTATTAGAAAATAAACAATAG 355

```

```

RESULT 7
BJ684174/c
LOCUS
DEFINITION BJ684174 HCEST library Haplochromis chilotes cDNA clone no90c12,
            mRNA sequence.
ACCESSION BJ684174
VERSION BJ684174.1 GI:46527295
KEYWORDS EST.
SOURCE Haplochromis chilotes
        ORGANISM Haplochromis chilotes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes;
            Labroidae; Cichlidae; Haplochromis.
REFERENCE 1 (bases 1 to 417)
AUTHORS Watanabe,M., Kobayashi,N., Shin-i,T., Kohara,Y. and Okada,N.
TITLE Orf sequences of cichlid in Lake Victoria are essentially same
JOURNAL Unpublished (2004)
COMMENT Contact: Tadasu Shin-i
        Center For Genetic Resource Information
        National Institute of Genetics
        1111 Yata, Mishima, Shizuoka 411-8540, Japan
        Tel: 81-559-81-6856
        Fax: 81-559-81-6855
        Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source           1..417
                     /organism="Haplochromis chilotes"
                     /mol_type="mRNA"
                     /db_xref="taxon:257977"
                     /clone="no90c12"
                     /tissue_type="jaw"
                     /dev_stage="varied"
                     /clone_lib="HCEST library"
ORIGIN
Query Match      100.0%; Score 101; DB 4; Length 417;
Best Local Similarity 100.0%; Pred. No. 8.3e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGTTATTGTCATGAGCGGATACATATTGTAATGTAATTTAGAAAAATAACAATAG 60
    |||||
Db 129 AGGTTATTGTCATGAGCGGATACATATTGTAATGTAATTTAGAAAAATAACAATAG 70

Qy 61 GGGTTCGGCGACATTTCCCGAAAAGTGCCACCTGACGTC 101
    |||||
Db 69 GGGTTCGGCGACATTTCCCGAAAAGTGCCACCTGACGTC 29

RESULT 8
CC819923/c
LOCUS
DEFINITION CC819923 Oxytricha plasmid UUGC10 library Sterkiella
            histriomuscorum genomic clone UUGC10006J13 R, genomic survey
            sequence.
ACCESSION CC819923
VERSION CC819923.1 GI:32900671
KEYWORDS GSS.
SOURCE Sterkiella histriomuscorum (Oxytricha trifallax)
        ORGANISM Sterkiella histriomuscorum
            Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
            Stichotrichida; Oxytrichidae; Sterkiella.
REFERENCE 1 (bases 1 to 491)
AUTHORS Dunn,D., Doak,T., Herrick,G. and Weiss,R.
TITLE Paired end reads from plasmid inserts of Oxytricha trifallax
JOURNAL macronuclear chromosomes
COMMENT Unpublished (2003)
        Contact: Robert B. Weiss
        University of Utah Genome Center
        University of Utah
        Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
        84112, USA
        Tel: 801 585 5606

```

```

Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Plate: 0006 row: J column: 13
Seq primer: CACACAGAAACACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 491.
FEATURES             Location/Qualifiers
     source           1..491
                     /organism="Sterkiella histriomuscorum"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:94289"
                     /clone="UUGC10006J13"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Oxytricha plasmid UUGC10 library"
                     /notes="Vector: FWD42nv; Purified macronuclear chromosomal
                     DNA from Oxytricha trifallax was blunt end-repaired with
                     T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
                     oligonucleotides were ligated to the blunt ends in high
                     molar excess. Vector DNA was prepared from a derivative of
                     PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
                     derivative of plasmid R1. The vector was ligated with
                     adaptors complementary to the insert adaptors and
                     purified. The sheared, adapted mouse DNA was annealed to
                     adapted vector DNA, and transformed into
                     chemically-competent E. Coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
ORIGIN
Query Match      100.0%; Score 101; DB 9; Length 491;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGTTATTGTCATGAGCGGATACATATTGTAATGTAATTTAGAAAAATAACAATAG 60
    |||||
Db 412 AGGTTATTGTCATGAGCGGATACATATTGTAATGTAATTTAGAAAAATAACAATAG 353

Qy 61 GGGTTCGGCGACATTTCCCGAAAAGTGCCACCTGACGTC 101
    |||||
Db 352 GGGTTCGGCGACATTTCCCGAAAAGTGCCACCTGACGTC 312

RESULT 9
BI805285
LOCUS
DEFINITION BI805285 Stem library from Oryza sativa (3-5 leaf stage) Oryza
            sativa cDNA clone S035A01, mRNA sequence.
ACCESSION BI805285
VERSION BI805285.1 GI:15852489
KEYWORDS EST.
SOURCE Oryza sativa
        ORGANISM Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 495)
AUTHORS Dong,H.T., Li,D.B., Zhuang,X.F., Dai,C.G., Sun,L.X., Pei,Y.X.,
            Wu,H.F., Jiang,Y.X., Yu,F.C., Gao,O.K. and Lou,Y.C.
TITLE A Gene Expression Screen in Oryza sativa
JOURNAL Unpublished (2001)
COMMENT Contact: Haitao Dong, Debao Li
        Bioinformatics and Gene Network Research Group
        Zhejiang University
        Kaixuan Road 268#, Hangzhou, Zhejiang, P.R.China
        Tel: 0086-571-8692051
        Fax: 0086-571-86961525
        Email: webmaster@estarray.org, URL: http://www.estarray.org
        Seq primer: M13 forward primer.
FEATURES             Location/Qualifiers
     source           1..495
                     /organism="Oryza sativa"
                     /mol_type="mRNA"
                     /db_xref="taxon:4530"
                     /clone="S035A01"

```

```
/tissue_type="Stem"
/dev_stage="3-5 leaf stage"
/clone_lib="Stem library from Oryza sativa (3-5 leaf
stage)"
/notes="Vector: pSport2"

ORIGIN
Query Match
Best Local Similarity 100.0%; Score 101; DB 4; Length 495;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATGTCATGACGCGATACATATTTGAATGATTTAGAAAAATAACAAATAG 60
Db |||||||
62 AGGGTTATGTCATGACGCGATACATATTTGAATGATTTAGAAAAATAACAAATAG 121
QY 61 GGGTTCCGCGACATTTCCCGAAAAGTGCCACCTGACGTC 101
Db |||||||
122 GGGTTCCGCGACATTTCCCGAAAAGTGCCACCTGACGTC 162

RESULT 10
CC818374/c
LOCUS
DEFINITION
100004807R Oxytricha plasmid UUGC100004B07 R, genomic survey
sequence.
ACCESSION
CC818374.1 GI:32897661
VERSION
CC818374
KEYWORDS
GSS.
SOURCE
Sterkiella histriomuscorum (Oxytricha trifallax)
ORGANISM
Sterkiella histriomuscorum
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
Stichotrichida; Oxytrichidae; Sterkiella.
REFERENCE
1 (bases 1 to 495)
Dunn,D., Doak,T., Herrick,G. and Weiss,R.
Paired end reads from plasmid inserts of Oxytricha trifallax
macronuclear chromosomes
Unpublished (2003)
JOURNAL
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Plate: 0004 row: B column: 07
Seq primer: CACACAGGAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 495.
FEATURES
Location/Qualifiers
1..495
/organism="Sterkiella histriomuscorum"
/mol_type="genomic DNA"
/db_xref="taxon:94289"
/clone="UUGC100004B07"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/notes="Vector: PWD42nv; Purified macronuclear chromosomal
DNA from Oxytricha trifallax was blunt end-repaired with
T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
oligonucleotides were ligated to the blunt ends in high
molar excess. Vector DNA was prepared from a derivative of
PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
derivative of plasmid R1. The vector was ligated with
adaptors complementary to the insert adaptors and
adapted vector DNA, and transformed into
chemically-competent E. Coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

ORIGIN
Query Match
Best Local Similarity 100.0%; Score 101; DB 9; Length 495;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATGTCATGACGCGATACATATTTGAATGATTTAGAAAAATAACAAATAG 60
Db |||||||
391 AGGGTTATGTCATGACGCGATACATATTTGAATGATTTAGAAAAATAACAAATAG 332

QY 61 GGGTTCCGCGACATTTCCCGAAAAGTGCCACCTGACGTC 101
Db |||||||
332 GGGTTCCGCGACATTTCCCGAAAAGTGCCACCTGACGTC 292

RESULT 11
CC818523/c
LOCUS
DEFINITION
100004113R Oxytricha plasmid UUGC100004L13 R, genomic survey
sequence.
ACCESSION
CC818523
VERSION
CC818523.1 GI:32897943
KEYWORDS
GSS.
SOURCE
Sterkiella histriomuscorum (Oxytricha trifallax)
ORGANISM
Sterkiella histriomuscorum
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
Stichotrichida; Oxytrichidae; Sterkiella.
REFERENCE
1 (bases 1 to 496)
Dunn,D., Doak,T., Herrick,G. and Weiss,R.
Paired end reads from plasmid inserts of Oxytricha trifallax
macronuclear chromosomes
Unpublished (2003)
JOURNAL
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Plate: 0004 row: L column: 13
Seq primer: CACACAGGAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 496.
FEATURES
Location/Qualifiers
1..496
/organism="Sterkiella histriomuscorum"
/mol_type="genomic DNA"
/db_xref="taxon:94289"
/clone="UUGC100004L13"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/notes="Vector: PWD42nv; Purified macronuclear chromosomal
DNA from Oxytricha trifallax was blunt end-repaired with
T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
oligonucleotides were ligated to the blunt ends in high
molar excess. Vector DNA was prepared from a derivative of
PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
derivative of plasmid R1. The vector was ligated with
adaptors complementary to the insert adaptors and
adapted vector DNA, and transformed into
chemically-competent E. Coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

ORIGIN
Query Match
Best Local Similarity 100.0%; Score 101; DB 9; Length 496;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATGTCATGACGCGATACATATTTGAATGATTTAGAAAAATAACAAATAG 60
Db |||||||
391 AGGGTTATGTCATGACGCGATACATATTTGAATGATTTAGAAAAATAACAAATAG 332

QY 61 GGGTTCCGCGACATTTCCCGAAAAGTGCCACCTGACGTC 101
Db |||||||
332 GGGTTCCGCGACATTTCCCGAAAAGTGCCACCTGACGTC 292
```

```

Db      331 GGGTTCCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 291
|||||
CC819854          503 bp      DNA      linear      GSS 17-JUL-2003
CC819854/c      100006N08R Oxytricha plasmid UUGC100006N08 R, genomic survey
LOCUS          histriomuscorum genomic clone UUGC100006N08 R, genomic survey
DEFINITION          sequence.

ACCESSION          CC819854
VERSION            CC819854.1  GI:32900533
KEYWORDS
SOURCE
ORGANISM          Sterkiella histriomuscorum (Oxytricha trifallax)
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
Stichotrichida; Oxytrichidae; Sterkiella.
REFERENCE          1 (bases 1 to 503)
AUTHORS            Dunn,D., Doak,T., Herrick,G. and Weiss,R.
TITLE              Paired end reads from plasmid inserts of Oxytricha trifallax
                  macronuclear chromosomes
JOURNAL            Unpublished (2003)
COMMENT            Contact: Robert B. Weiss
                  University of Utah Genome Center
                  Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
                  84112, USA
                  Tel: 801 585 5606
                  Fax: 801 585 7177
                  Email: ddunn@genetics.utah.edu
                  Plate: 0006 row: N column: 08
                  Seq primer: CACACAGGAACAGCTATGACC
                  Class: plasmid ends
                  High quality sequence stop: 503.
FEATURES          source
                  1..503
                   /organism="Sterkiella histriomuscorum"
                   /mol_type="genomic DNA"
                   /db_xref="taxon:94289"
                   /clone="UUGC100006N08"
                   /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
                   /note="Vector: PWD42nv; Purified macronuclear chromosomal
                   DNA from Oxytricha trifallax was blunt end-repaired with
                   T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
                   oligonucleotides were ligated to the blunt ends in high
                   molar excess. Vector DNA was prepared from a derivative of
                   pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
                   derivative of plasmid R1. The vector was ligated with
                   adaptors complementary to the insert adaptors and
                   purified. The sheared, adaptor mouse DNA was annealed to
                   adapted vector DNA, and transformed into
                   chemically-competent E. Coli XL10-Gold (Stratagene) cells
                   and selected for ampicillin resistance."
ORIGIN
Query Match          100.0%; Score 101; DB 9; Length 503;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAACAAATAG 60
|||||
Db      410 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAACAAATAG 351
|||||

Qy      61 GGGTTCCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
|||||
Db      350 GGGTTCCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 310
|||||

RESULT 13
CC817752/c      CC817752          515 bp      DNA      linear      GSS 17-JUL-2003
LOCUS          100003C16R Oxytricha plasmid UUGC10 library Sterkiella
DEFINITION          Sterkiella histriomuscorum
                  Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
                  Sterkiella histriomuscorum (Oxytricha trifallax)
                  sequence.
ACCESSION          CC817752
VERSION            CC817752.1  GI:32896415
KEYWORDS
SOURCE
ORGANISM          Sterkiella histriomuscorum (Oxytricha trifallax)
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
Sterkiella histriomuscorum
sequence.
ACCSSION          CC817128
VERSION            CC817128.1  GI:32896415
KEYWORDS
SOURCE
ORGANISM          Sterkiella histriomuscorum (Oxytricha trifallax)
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
Sterkiella histriomuscorum
sequence.
ACCSSION          CC817752
VERSION            CC817752.1  GI:32897039
KEYWORDS
SOURCE
ORGANISM          Sterkiella histriomuscorum (Oxytricha trifallax)
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
Stichotrichida; Oxytrichidae; Sterkiella.
REFERENCE          1 (bases 1 to 515)
AUTHORS            Dunn,D., Doak,T., Herrick,G. and Weiss,R.
TITLE              Paired end reads from plasmid inserts of Oxytricha trifallax
                  macronuclear chromosomes
JOURNAL            Unpublished (2003)
COMMENT            Contact: Robert B. Weiss
                  University of Utah Genome Center
                  Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
                  84112, USA
                  Tel: 801 585 5606
                  Fax: 801 585 7177
                  Email: ddunn@genetics.utah.edu
                  Plate: 0003 row: C column: 16
                  Seq primer: CACACAGGAACAGCTATGACC
                  Class: plasmid ends
                  High quality sequence stop: 515.
FEATURES          source
                  1..515
                   /organism="Sterkiella histriomuscorum"
                   /mol_type="genomic DNA"
                   /db_xref="taxon:94289"
                   /clone="UUGC100003C16"
                   /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
                   /note="Vector: PWD42nv; Purified macronuclear chromosomal
                   DNA from Oxytricha trifallax was blunt end-repaired with
                   T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
                   oligonucleotides were ligated to the blunt ends in high
                   molar excess. Vector DNA was prepared from a derivative of
                   pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
                   derivative of plasmid R1. The vector was ligated with
                   adaptors complementary to the insert adaptors and
                   purified. The sheared, adaptor mouse DNA was annealed to
                   adapted vector DNA, and transformed into
                   chemically-competent E. Coli XL10-Gold (Stratagene) cells
                   and selected for ampicillin resistance."
ORIGIN
Query Match          100.0%; Score 101; DB 9; Length 515;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAACAAATAG 60
|||||
Db      412 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAACAAATAG 353
|||||

Qy      61 GGGTTCCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
|||||
Db      352 GGGTTCCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 312
|||||

RESULT 14
CC817128/c      CC817128          518 bp      DNA      linear      GSS 17-JUL-2003
LOCUS          100002D21R Oxytricha plasmid UUGC10 library Sterkiella
DEFINITION          histriomuscorum genomic clone UUGC100002D21 R, genomic survey
                  sequence.
ACCESSION          CC817128
VERSION            CC817128.1  GI:32896415
KEYWORDS
SOURCE
ORGANISM          Sterkiella histriomuscorum (Oxytricha trifallax)
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
Sterkiella histriomuscorum
sequence.

```

```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Stichotrichida, Oxytrichidae; Sterkiella.
1 (bases 1 to 518)
Dunn,D., Doak,T., Herrick,G. and Weiss,R.
Paired end reads from plasmid inserts of Oxytricha trifallax
macronuclear chromosomes
Unpublished (2003)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Plate: 0002 row: D column: 21
Seq primer: CACACAGGAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 518.

FEATURES
source
1..518
/organism="Sterkiella histriomuscorum"
/mol_type="genomic DNA"
/db_xref="taxon:94289"
/clone="UUGC100002D21"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Oxytricha plasmid UUGC10 library"
/note="Vector: PWD42nv; Purified macronuclear chromosomal
DNA from Oxytricha trifallax was blunt end-repaired with
T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
oligonucleotides were ligated to the blunt ends in high
molar excess. Vector DNA was prepared from a derivative of
pWD42 (G14732114[gblAF129072.1]), a copy-number inducible
derivative of plasmid R1. The vector was ligated with
adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. Coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

ORIGIN
Query Match 100.0%; Score 101; DB 9; Length 518;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTAGAAAAATAACAATAG 60
Db 410 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTAGAAAAATAACAATAG 351

Qy 61 GGGTTCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 101
Db 350 GGGTTCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 310

RESULT 15
CC817162/c
LOCUS
DEFINITION
100002J19R Oxytricha plasmid UUGC10 library Sterkiella
histriomuscorum genomic clone UUGC100002J19 R, genomic survey
sequence.
CC817162
CC817162.1 GI:32896449
GSS.
Sterkiella histriomuscorum (Oxytricha trifallax)
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
Stichotrichida; Oxytrichidae; Sterkiella.
1 (bases 1 to 519)
Dunn,D., Doak,T., Herrick,G. and Weiss,R.
Paired end reads from plasmid inserts of Oxytricha trifallax
macronuclear chromosomes
Unpublished (2003)
Contact: Robert B. Weiss
University of Utah Genome Center

```

```

University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Plate: 0002 row: J column: 19
Seq primer: CACACAGGAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 519.

FEATURES
Location/Qualifiers
1..519
/organism="Sterkiella histriomuscorum"
/mol_type="genomic DNA"
/db_xref="taxon:94289"
/clone="UUGC100002J19"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Oxytricha plasmid UUGC10 library"
/note="Vector: PWD42nv; Purified macronuclear chromosomal
DNA from Oxytricha trifallax was blunt end-repaired with
T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
oligonucleotides were ligated to the blunt ends in high
molar excess. Vector DNA was prepared from a derivative of
pWD42 (G14732114[gblAF129072.1]), a copy-number inducible
derivative of plasmid R1. The vector was ligated with
adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. Coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

ORIGIN
Query Match 100.0%; Score 101; DB 9; Length 519;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTAGAAAAATAACAATAG 60
Db 416 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTAGAAAAATAACAATAG 357

Qy 61 GGGTTCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 101
Db 356 GGGTTCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 316

Search completed: July 14, 2005, 23:23:05
Job time : 962.667 secs

```

THIS PAGE BLANK (USPTO)

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	100	100.0	3853	6	AR098190	Sequence
2	100	100.0	3853	6	AR207832	Sequence
3	100	100.0	3853	6	BD009729	Tissue sp
4	100	100.0	3986	12	PCDNA32B0	X90639 Cloning vec
5	100	100.0	4026	6	AR098191	Sequence
6	100	100.0	4026	6	AR207833	Sequence
7	100	100.0	4026	6	BD009730	Tissue sp
8	100	100.0	4249	6	AR098192	Sequence
9	100	100.0	4249	6	AR207834	Sequence
10	100	100.0	4249	6	BD009731	Tissue sp
11	100	100.0	4341	6	A38214	Sequence 58
12	100	100.0	4341	6	AX286570	Sequence
13	100	100.0	4597	6	AX060344	Sequence
14	100	100.0	4840	6	AX133940	Sequence
15	100	100.0	5053	6	BD238492	Sequence
16	100	100.0	5070	6	BD238492	Expressio
17	100	100.0	5082	6	AX234391	Sequence
18	100	100.0	5082	6	A91754	Sequence 10
19	100	100.0	5082	6	BD085110	Vertebra
20	100	100.0	5162	6	AX951626	Sequence

```
Unclassified.
REFERENCE 1 (bases 1 to 3853)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion proteins
JOURNAL Patent: US 6379927-A 5 30-APR-2002;
FEATURES Location/Qualifiers
source 1..3853
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 3853;
Best Local Similarity 100.0%; Pred. No. 9.4e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
Db 1 GACGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
Qy 61 CGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
Db 61 CGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

RESULT 3
BD009729
LOCUS BD009729 3853 bp DNA linear PAT 31-JAN-2002
DEFINITION Tissue specific expression of retinoblastoma protein.
ACCESSION BD009729
VERSION BD009729.1 GI:18638102
KEYWORDS JP 2001503638-A/3.
SOURCE unidentified
ORGANISM unidentified
unclassified.
REFERENCE 1 (bases 1 to 3853)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Tissue specific expression of retinoblastoma protein
JOURNAL Patent: JP 2001503638-A 3 21-MAR-2001;
COMMENT CANUI INC
OS Unidentified
PN JP 2001503638-A/3
PD 21-MAR-2001
PF 13-NOV-1997 JP 1998522958
PR 15-NOV-1996 US 08/751517,14-FEB-1997 US 08/801092 PI
C07H21/04,C07K5/00,A61K38/00,A61K35/12
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT CDS 209..862.
FEATURES source
1..3853
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 3853;
Best Local Similarity 100.0%; Pred. No. 9.4e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
Db 1 GACGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
Qy 61 CGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
Db 61 CGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

PCDNA3ZEO 3986 bp DNA linear SYN 16-AUG-1995

RESULT 4
PCDNA3ZEO
LOCUS PCDNA3ZEO 3986 bp DNA linear SYN 16-AUG-1995
DEFINITION Cloning vector pCDNA3ZEO DNA.
ACCESSION X90639
VERSION X90639.1 GI:949972
KEYWORDS cloning vector; expression vector; multiple cloning site; Plasmid.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Peters,H., Hunthausen,T., Kroenke,M. and Marget,M.
TITLE A new small sized high-level eukaryotic expression vector
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3986)
AUTHORS Peters,H.
TITLE Direct Submission
JOURNAL Michaelisstr.5, D- 24105 Kiel, FRG
COMMENT Related sequences: M21295 and K03104.
FEATURES Location/Qualifiers
source 1..3986
/organism="synthetic construct"
/mol_type="other DNA"
/db_xref="taxon:32630"
/plasmid="pCDNA3ZEO"
misc_feature 1..2125
/notes="cloning vector (pCDNA3) (Invitrogen)"
misc_feature 889..994
/notes="multiple cloning site (MCS)"
misc_feature 2126..2796
/notes="cloning vector (PzeoSV) (Invitrogen)"
misc_feature 2797..3986
/notes="cloning vector (pCDNA3)"

ORIGIN
Query Match 100.0%; Score 100; DB 12; Length 3986;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
Db 1 GACGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
Qy 61 CGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
Db 61 CGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

RESULT 5
AR098191
LOCUS AR098191 4026 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 19 from patent US 6074850.
ACCESSION AR098191
VERSION AR098191.1 GI:12807448
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 4026)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion polypeptides
JOURNAL Patent: US 6074850-A 19 13-JUN-2000;
FEATURES Location/Qualifiers
source 1..4026
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4026;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
Db 1 GACGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
```



```
QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGCTGTT 100
    |||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGCTGTT 100

RESULT 6
AR207833
LOCUS AR207833 4026 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 19 from patent US 6379927.
ACCESSION AR207833
VERSION AR207833.1 GI:21507689
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 4026)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion proteins
JOURNAL Patent: US 6379927-A 19 30-APR-2002;
FEATURES Location/Qualifiers
source 1..4026
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4026;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
    |||||
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60

RESULT 7
AR207833
LOCUS AR207833 4026 bp DNA linear PAT 31-JAN-2002
DEFINITION Tissue specific expression of retinoblastoma protein.
ACCESSION BD009730
VERSION BD009730.1 GI:18638103
KEYWORDS JP 2001503638-A/4.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 4026)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Tissue specific expression of retinoblastoma protein
JOURNAL Patent: JP 2001503638-A 4 21-MAR-2001;
COMMENT CANJ INC
OS Unidentified
PN JP 2001503638-A/4
PD 21-MAR-2001
PF 13-NOV-1997 JP 1998522958
PR 15-NOV-1996 US 08/751517,14-FEB-1997 US 08/801092 PI
COU7H21/04,C07K5/00,A61K38/00,A61K35/12
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1..4026
/mol_type="Unidentified".
FEATURES Location/Qualifiers
source 1..4026
/mol_type="unidentified"
/db_xref="taxon:32644"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4026;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
    |||||
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGCTGTT 100
    |||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGCTGTT 100

RESULT 8
AR098192
LOCUS AR098192 4249 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 33 from patent US 6074850.
ACCESSION AR098192
VERSION AR098192.1 GI:12807449
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 4249)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion polypeptides
JOURNAL Patent: US 6074850-A 33 13-JUN-2000;
FEATURES Location/Qualifiers
source 1..4249
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4249;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
    |||||
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGCTGTT 100
    |||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGCTGTT 100

RESULT 9
AR207834
LOCUS AR207834 4249 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 33 from patent US 6379927.
ACCESSION AR207834
VERSION AR207834.1 GI:21507691
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 4249)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion proteins
JOURNAL Patent: US 6379927-A 33 30-APR-2002;
FEATURES Location/Qualifiers
source 1..4249
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4249;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
    |||||
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGCTGTT 100
    |||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGCTGTT 100

RESULT 10
AR207834
LOCUS AR207834 4249 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 33 from patent US 6379927.
ACCESSION AR207834
VERSION AR207834.1 GI:21507691
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 4249)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion proteins
JOURNAL Patent: US 6379927-A 33 30-APR-2002;
FEATURES Location/Qualifiers
source 1..4249
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4249;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
    |||||
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
```

```
Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

RESULT 10
BD009731
LOCUS BD009731 4249 bp DNA linear PAT 31-JAN-2002
DEFINITION Tissue specific expression of retinoblastoma protein.
ACCESSION BD009731
VERSION BD009731.1 GI:18638104
KEYWORDS JP 2001503638-A/5.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 4249)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Tissue specific expression of retinoblastoma protein
JOURNAL Patent: JP 2001503638-A 5 21-MAR-2001;
CANJ I INC
COMMENT OS Unidentified
PN JP 2001503638-A/5
PD 21-MAR-2001
PF 13-NOV-1997 JP 1998522958
PR 15-NOV-1996 US 08/751517,14-FEB-1997 US 08/801092 PT
DOUGLAS ANTELMAN,RICHARD J GREGORY,KENNETH N WILLS PC
C07H21/04,C07K5/00,A61K38/00,A61K35/12
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1..4249
FT Location/Qualifiers
FEATURES source
1..4249
/organism="Unidentified".
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4249;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAACTGCTCTGATG 60
Db 1 GACGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAACTGCTCTGATG 60
Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

RESULT 11
A38214
LOCUS A38214 4341 bp DNA linear PAT 05-MAR-1997
DEFINITION Sequence 58 from Patent WO9408008.
ACCESSION A38214
VERSION A38214.1 GI:2294819
KEYWORDS unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 4341)
AUTHORS Hawkins,R.E., Russell,S.J., Stevenson,F.K. and Winter,G.P.
TITLE IMPROVEMENTS IN OR RELATING TO IMMUNE RESPONSE MODIFICATION
JOURNAL Patent: WO 9408008-A 58 14-APR-1994;
MEDICAL RES COUNCIL (GB)
COMMENT Other publication CA 2145064 940414
Other publication AU 4832493 940426
Other publication JP 8501699T 960227.
FEATURES source
1..4341
Location/Qualifiers

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

RESULT 12
AX286570
LOCUS AX286570 4341 bp DNA linear PAT 21-NOV-2001
DEFINITION Sequence 1 from Patent WO0179510.
ACCESSION AX286570
VERSION AX286570.1 GI:17048664
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Rice,J.H. and Stevenson,F.M.
TITLE Materials and methods relating to immune responses to fusion
JOURNAL Patent: WO 0179510-A 1 25-OCT-2001;
Cancer Research Ventures Limited (GB)
FEATURES source
1..4341
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/notes="Vector pVAC1"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4341;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAACTGCTCTGATG 60
Db 1 GACGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAACTGCTCTGATG 60
Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

RESULT 13
AX060344
LOCUS AX060344 4597 bp DNA linear PAT 22-JAN-2001
DEFINITION Sequence 3 from Patent WO0078358.
ACCESSION AX060344
VERSION AX060344.1 GI:12405832
KEYWORDS synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Chen,W.
TITLE Hyaluronic acid microspheres for sustained gene transfer
JOURNAL Patent: WO 0078358-A 3 28-DEC-2000;
The Collaborative Group, Ltd. (US)
FEATURES source
1..4597
Location/Qualifiers
/organism="synthetic construct"
```

```
ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4341;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAACTGCTCTGATG 60
Db 1 GACGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAACTGCTCTGATG 60
Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

RESULT 12
AX286570
LOCUS AX286570 4341 bp DNA linear PAT 21-NOV-2001
DEFINITION Sequence 1 from Patent WO0179510.
ACCESSION AX286570
VERSION AX286570.1 GI:17048664
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Rice,J.H. and Stevenson,F.M.
TITLE Materials and methods relating to immune responses to fusion
JOURNAL Patent: WO 0179510-A 1 25-OCT-2001;
Cancer Research Ventures Limited (GB)
FEATURES source
1..4341
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/notes="Vector pVAC1"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4341;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAACTGCTCTGATG 60
Db 1 GACGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAACTGCTCTGATG 60
Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

RESULT 13
AX060344
LOCUS AX060344 4597 bp DNA linear PAT 22-JAN-2001
DEFINITION Sequence 3 from Patent WO0078358.
ACCESSION AX060344
VERSION AX060344.1 GI:12405832
KEYWORDS synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Chen,W.
TITLE Hyaluronic acid microspheres for sustained gene transfer
JOURNAL Patent: WO 0078358-A 3 28-DEC-2000;
The Collaborative Group, Ltd. (US)
FEATURES source
1..4597
Location/Qualifiers
/organism="synthetic construct"
```

```
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="pCDNA3.1/GS vector by Invitrogen Corporation"

ORIGIN
Query Match          100.0%; Score 100; DB 6; Length 4597;
Best Local Similarity 100.0%; Pred. No. 9.2e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
    |||||||
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
    |||||||

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
    |||||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
    |||||||

RESULT 14
AXI33940
LOCUS AXI33940 4840 bp DNA linear PAT 15-MAY-2001
DEFINITION Sequence 1 from Patent WO0119853.
ACCESSION AXI33940
VERSION AXI33940.1 GI:14139881
KEYWORDS
SOURCE
ORGANISM
synthetic construct
other sequences; artificial sequences.
REFERENCE
AUTHORS Hollander,A.P., Barker,M.D. and Kafienah,W.C.
TITLE Cell transfection
JOURNAL Patent: WO 0119853-A 1 22-MAR-2001;
THE UNIVERSITY OF SHEFFIELD (GB)
FEATURES
Location/Qualifiers
source 1. 4840
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="This sequence is artificial and is based on well
established commercially available vectors that are cited
with their vendor within the patent application"

ORIGIN
Query Match          100.0%; Score 100; DB 6; Length 4840;
Best Local Similarity 100.0%; Pred. No. 9.2e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
    |||||||
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
    |||||||

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
    |||||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
    |||||||

RESULT 15
BD238492
LOCUS BD238492 5053 bp DNA linear PAT 17-JUL-2003
DEFINITION Expression vectors for stimulating an immune response and methods
of using the same.
ACCESSION BD238492
VERSION BD238492.1 GI:33048262
KEYWORDS JP 2002520000-A/18.
SOURCE
ORGANISM
synthetic construct
other sequences; artificial sequences.
REFERENCE
AUTHORS Fikes,J.D., Hermanson,G.G., Sette,A., Ishioka,G.Y., Livingston,B.
and Chesnut,R.W.
TITLE Expression vectors for stimulating an immune response and methods
of using the same
JOURNAL Patent: JP 2002520000-A 18 09-JUL-2002;
```

```
EPIMMUNE INC
OS Artificial Sequence
PN JP 2002520000-A/18
PD 09-JUL-2002
PF 13-MAY-1999 JP 2000548449
PR 13-MAY-1998 US 09/078904,15-MAY-1998 US 60/085751 PI
JOHN D FIKES,GARY G HERMANSON,ALESSANDRO
SETTE,GLENN Y ISHIOKA,
PI BRIAN LIVINGSTON,ROBERT W CHESNUT
PC C12N15/09,A61K31/711,A61K39/00,A61K39/12,A61K39/21,A61K39/29,
PC A61P31/14,
PC A61P31/16,A61P31/20,A61P37/02,C12N15/00
CC vector pEP2
FH Key Location/Qualifiers
FT source 1..5053
/organism="Artificial Sequence".
FEATURES
source 1..5053
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

ORIGIN
Query Match          100.0%; Score 100; DB 6; Length 5053;
Best Local Similarity 100.0%; Pred. No. 9.2e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
    |||||||
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATG 60
    |||||||

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
    |||||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
    |||||||

Search completed: July 14, 2005, 14:03:32
Job time : 749.127 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright. (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:35:42 ; Search time 140.988 Seconds
(without alignments)
4198.742 Million cell updates/sec

Title: US-09-482-682-47_COPY_1_100

Perfect score: 100

Sequence: 1 gacggatcggagatctccc.....ctgtccctcgttgtgtgtt 100

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1: Geneseq1980s:*

2: Geneseq1990s:*

3: Geneseq2000s:*

4: Geneseq2001as:*

5: Geneseq2001bs:*

6: Geneseq2002as:*

7: Geneseq2002bs:*

8: Geneseq2003as:*

9: Geneseq2003bs:*

10: Geneseq2003cs:*

11: Geneseq2003ds:*

12: Geneseq2004as:*

13: Geneseq2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100	100.0	1506	12 ADM41035	Adm41035 Fungus nu
2	100	100.0	1600	2 ADH11349	Adh11349 Vertebrat
3	100	100.0	1782	12 ADM41037	Adm41037 Cytomegal
4	100	100.0	2241	12 ADM41034	Adm41034 Human nuc
5	100	100.0	2294	12 ADM41036	Adm41036 Cytomegal
6	100	100.0	3853	2 AAV40006	Aav40006 Plasmid p
7	100	100.0	4026	2 AAV40007	Aav40007 Plasmid p
8	100	100.0	4249	2 AAV63466	Aav63466 Plasmid p
9	100	100.0	4341	2 AAG62391	Aag62391 Vector pV
10	100	100.0	4341	6 AAS17704	Aas17704 Vector pV
11	100	100.0	4341	6 AEN83143	Aen83143 Plasmid p
12	100	100.0	4597	4 AAF24901	Aaf24901 Nucleotid
13	100	100.0	4639	6 AAD39652	Aad39652 Human sma
14	100	100.0	4840	4 AAF83146	Aaf83146 Complete
15	100	100.0	5015	10 ADB33528	Adb33528 Expressio
16	100	100.0	5053	3 AAZ38633	Aaz38633 pE2 expr
17	100	100.0	5070	4 AAS12839	Aas12839 DNA seque
18	100	100.0	5082	2 ADH11417	Adh11417 Plasmid p
19	100	100.0	5162	10 ADF10526	Adf10526 Plasmid p
20	100	100.0	5162	10 ACC44637	Acc44637 Murine rD

ALIGNMENTS

RESULT 1

ADMA41035

ID ADM41035 standard; DNA; 1506 BP.

XX AC ADM41035;

XX DT 17-JUN-2004 (first entry)

XX DE Fungus nucleotide sequence SEQ ID NO:3.

XX KW engrafting foreign replacement cell; implanting foreign replacement cell;

XX KW growth; differentiation; drug development; vaccine development;

XX KW tissue transplantation; human disease study; fungus; gene; ds.

XX OS Unidentified.

XX FN WO2004027029-A2.

XX PD 01-APR-2004.

XX PF 17-SEP-2003; 2003WO-US029251.

XX PR 19-SEP-2002; 2002US-0411790P.

XX PA (XIME-) XIMEREX INC.

XX PI Beechornor WE, Sosa CE, Thompson SC;

XX XX WPI; 2004-295402/27.

XX DR Engrafting foreign replacement cells within a fetal non-human mammal,

XX XX useful in producing chimeric mammals, comprises selectively destroying

XX PT native cells in a tissue of a fetal non-human mammal host.

XX PT Disclosure; SEQ ID NO 3; 48pp; English.

XX XX The present invention describes a method for engrafting foreign

XX PS replacement cells within a foetal non-human mammal, which comprises

XX CC selectively destroying native cells in a tissue of a foetal non-human

XX CC mammal host, where the number of maternal cells of the same tissue is not

XX CC substantially reduced, and implanting foreign replacement cells in the

XX CC tissue of the fetal non-human mammal host, where the foreign replacement

XX CC cells replace destroyed cells of the tissue. The method is useful for

AdS75099 Plasmid p
Acc44692 Plasmid p
Abv77540 Plasmid p
Abv77538 Plasmid p
Abv77549 Plasmid p
Adi34681 Nucleotid
Abv77539 Plasmid p
Adh11233 Vertebrat
Abn86685 Nucleotid
Ade21866 Plasmid v
Ado05277 pcDNA3 pl
Aaz89476 Transgeni
Aasi18619 Renilla l
Adi53540 Vector pc
Adn36314 Plasmid p
Abi58494 Recombina
Abi58493 Recombina
Abk88868 Topoisome
Ade83791 Plasmid p
Ado06720 Recombina
Abi58489 Recombina
Abi58490 Recombina
Aag88310 Plasmid p
Aai66195 Human FSH
Adk40237 DNA encod

CC facilitating growth and differentiation of foreign cells within a
CC mammalian host, and for producing chimeric mammals that can be used to
CC develop new drugs and vaccine, factors, drugs and tissues for
CC transplantation, also useful to study human diseases. The present
CC sequence represents a nucleotide sequence given in the Sequence Listing
CC of the present invention but not mentioned further within the
CC specification.

XX
SQ Sequence 1506 BP; 454 A; 277 C; 361 G; 414 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 12; Length 1506;
Best Local Similarity 100.0%; Pred. No. 4e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCGATCCCTATGCTGCTCAGTACAAATCTGCTCTGATG 60
|||||
Db 1 GACGGATCGGAGATCTCCGATCCCTATGCTGCTCAGTACAAATCTGCTCTGATG 60
|||||

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
|||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
|||||

RESULT 2
ADH11349
ID ADH11349 standard; DNA; 1600 BP.

XX
AC ADH11349;
XX
DT 11-MAR-2004 (first entry)
XX

DE Vertebrate UNC-53 protein homologue related nucleotide sequence.

XX
KW UNC-53 vertebrate protein homologue; UNC-53; Caenorhabditis elegans;
KW cell shape regulator; cell motility regulator; cell migration;
KW cell behaviour regulator; phenotype; signal transduction pathway;
KW signal transducing protein; signal integrator protein;
KW neuronal regeneration; revascularisation; wound healing;
KW chronic neurodegenerative disease; acute traumatic injury;
KW fibrotic disease; gene; ds.

XX
OS Unidentified.

XX
FN WO9824810-A2.
XX
PD 11-JUN-1998.
XX
PF 03-DEC-1997; 97WO-EP006956.
XX
PR 04-DEC-1996; 96GB-00025283.
XX
PA (JANC) JANSSEN PHARM NV.
XX
PI Platteeuw CJ, Buesa Arjol CM, Deraeymaeker M, Verhasselt P;
PI Pujol NJR, Maertens LUS, Luyten W, Geerts H, Vandekerckhove JS;
PI Geysen J, Bogaert TAOE;
XX
XX WPI; 1998-362411/31.
DR P-PSDB; ADH11350.
XX
PT Vertebrate homologues of C. elegans UNC-53 protein - useful for, e.g.
PT promoting neuronal regeneration, treating chronic neuro-degenerative
PT diseases or acute traumatic injuries.

PS Disclosure; Page 410-411; 479pp; English.

XX
CC The present invention describes a vertebrate protein homologue of an UNC-
CC 53 protein of Caenorhabditis elegans or a functional equivalent,
CC derivative or bioprecursor of UNC-53. Also described: (1) a cDNA sequence
CC encoding a vertebrate homologue of the C. elegans UNC-53 protein; (2) a
CC nucleic acid which hybridises to the cDNA of (1); (3) vector comprising
CC the cDNA as in (1); (4) a host cell containing the vector as in (3); (5)
CC a transgenic cell, tissue or animal comprising the vector as in (3); (6)

CC a compound identified as an enhancer or inhibitor of the regulation of
CC cell shape, motility, or the direction of cell migration for use as a
CC therapeutic; (7) a method for determination of whether a protein is an
CC inhibitor or enhancer of regulation of cell behaviour, growth, shape or
CC motility or the direction of migration by contacting a host cell
CC expressing a homologue of UNC-53 and determining a change of phenotype;
CC (8) a method for identification of vertebrate homologues of C. elegans
CC unc-53 gene by hybridising a probe of 15-50 bp of the unc-53 sequence to
CC a DNA library; and (9) a method for identification of a protein which is
CC active in the signal transduction pathway of a cell of which a vertebrate
CC homologue of UNC-53 is a component comprising: (i) contacting an extract
CC of a cell with an antibody to the UNC-53 homologue; (ii) identifying an
CC antibody/homologue complex; and (iii) analysing such a complex to
CC identify any non-antibody protein bound to the complex. UNC-53 is a
CC signal transducing or signal integrator protein involved in controlling
CC directionality of cell migration and cell shape in C. elegans. Vertebrate
CC homologues of UNC-53 can be used to promote neuronal regeneration,
CC revascularisation or wound healing, to treat chronic neurodegenerative
CC diseases or acute traumatic injuries or fibrotic diseases. The present
CC sequence is used in the exemplification of the present invention.

XX
SQ Sequence 1600 BP; 397 A; 409 C; 398 G; 396 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 2; Length 1600;
Best Local Similarity 100.0%; Pred. No. 4.1e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCGATCCCTATGCTGCTCAGTACAAATCTGCTCTGATG 60
|||||
Db 1 GACGGATCGGAGATCTCCGATCCCTATGCTGCTCAGTACAAATCTGCTCTGATG 60
|||||

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
|||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
|||||

RESULT 3
ADH41037
ID ADH41037 standard; DNA; 1782 BP.

XX
AC ADH41037;
XX
DT 17-JUN-2004 (first entry)
XX
DE Cytomegalovirus nucleotide sequence SEQ ID NO:5.
XX
KW engrafting foreign replacement cell; implanting foreign replacement cell;
KW growth; differentiation; drug development; vaccine development;
KW tissue transplantation; human disease study; cytomegalovirus; gene; ds.

XX
OS Cytomegalovirus.

XX
FN WO2004027029-A2.
XX
PD 01-APR-2004.
XX
PF 17-SEP-2003; 2003WO-US029251.
XX
PR 19-SEP-2002; 2002US-0411790P.
XX
PA (XIME-) XIMEREX INC.
XX
PI Beschoner WE, Sosa CE, Thompson SC;
XX
XX WPI; 2004-295402/27.
XX
PT Engrafting foreign replacement cells within a fetal non-human mammal,
PT useful in producing chimeric mammals, comprises selectively destroying
PT native cells in a tissue of a fetal non-human mammal host.

PS Disclosure; SEQ ID NO 5; 48pp; English.

XX
CC The present invention describes a method for engrafting foreign

CC replacement cells within a foetal non-human mammal, which comprises
CC selectively destroying native cells in a tissue of a foetal non-human
CC mammal host, where the number of maternal cells of the same tissue is not
CC substantially reduced, and implanting foreign replacement cells in the
CC tissue of the foetal non-human mammal host, where the foreign replacement
CC cells replace destroyed cells of the tissue. The method is useful for
CC facilitating growth and differentiation of foreign cells within a
CC mammalian host, and for producing chimeric mammals that can be used to
CC develop new drugs and vaccine, factors, drugs and tissues for
CC transplantation, also useful to study human diseases. The present
CC sequence represents a nucleotide sequence given in the Sequence Listing
CC of the present invention but not mentioned further within the
CC specification.

XX SQ Sequence 1782 BP; 458 A; 399 C; 451 G; 474 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 12; Length 1782;
Best Local Similarity 100.0%; Pred. No. 4.2e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

RESULT 4

ADM41034
ID ADM41034 standard; DNA; 2241 BP.

XX AC ADM41034;

XX DT 17-JUN-2004 (first entry)

XX DE Human nucleotide sequence SEQ ID NO:2.

XX KW engrafting foreign replacement cell; implanting foreign replacement cell;
XX KW growth; differentiation; drug development; vaccine development;
XX KW tissue transplantation; human disease study; human; gene; ds.

XX OS Homo sapiens.

XX PN WO2004027029-A2.

XX PD 01-APR-2004.

XX PF 17-SEP-2003; 2003WO-US029251.

XX PR 19-SEP-2002; 2002US-0411790P.

XX FA (XIME-) XIMEREX INC.

XX PI Beschornier WE, Sosa CE, Thompson SC;

XX DR WPI; 2004-295402/27.

XX PT Engrafting foreign replacement cells within a foetal non-human mammal,
XX PT useful in producing chimeric mammals, comprises selectively destroying
XX PT native cells in a tissue of a foetal non-human mammal host.

XX PS Disclosure; SEQ ID NO 2; 48pp; English.

XX CC The present invention describes a method for engrafting foreign
XX CC replacement cells within a foetal non-human mammal, which comprises
XX CC selectively destroying native cells in a tissue of a foetal non-human
XX CC mammal host, where the number of maternal cells of the same tissue is not
XX CC substantially reduced, and implanting foreign replacement cells in the
XX CC tissue of the foetal non-human mammal host, where the foreign replacement
XX CC cells replace destroyed cells of the tissue. The method is useful for
XX CC facilitating growth and differentiation of foreign cells within a

CC mammalian host, and for producing chimeric mammals that can be used to
CC develop new drugs and vaccine, factors, drugs and tissues for
CC transplantation, also useful to study human diseases. The present
CC sequence represents a nucleotide sequence given in the Sequence Listing
CC of the present invention but not mentioned further within the
CC specification.

XX SQ Sequence 2241 BP; 498 A; 630 C; 609 G; 504 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 12; Length 2241;
Best Local Similarity 100.0%; Pred. No. 4.5e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

RESULT 5

ADM41036
ID ADM41036 standard; DNA; 2294 BP.

XX AC ADM41036;

XX DT 17-JUN-2004 (first entry)

XX DE Cytomegalovirus nucleotide sequence SEQ ID NO:4.

XX KW engrafting foreign replacement cell; implanting foreign replacement cell;
XX KW growth; differentiation; drug development; vaccine development;
XX KW tissue transplantation; human disease study; cytomegalovirus; gene; ds.

XX OS Cytomegalovirus.

XX PN WO2004027029-A2.

XX PD 01-APR-2004.

XX PF 17-SEP-2003; 2003WO-US029251.

XX PR 19-SEP-2002; 2002US-0411790P.

XX FA (XIME-) XIMEREX INC.

XX PI Beschornier WE, Sosa CE, Thompson SC;

XX DR WPI; 2004-295402/27.

XX PT Engrafting foreign replacement cells within a foetal non-human mammal,
XX PT useful in producing chimeric mammals, comprises selectively destroying
XX PT native cells in a tissue of a foetal non-human mammal host.

XX PS Disclosure; SEQ ID NO 4; 48pp; English.

XX CC The present invention describes a method for engrafting foreign
XX CC replacement cells within a foetal non-human mammal, which comprises
XX CC selectively destroying native cells in a tissue of a foetal non-human
XX CC mammal host, where the number of maternal cells of the same tissue is not
XX CC substantially reduced, and implanting foreign replacement cells in the
XX CC tissue of the foetal non-human mammal host, where the foreign replacement
XX CC cells replace destroyed cells of the tissue. The method is useful for
XX CC facilitating growth and differentiation of foreign cells within a
XX CC mammalian host, and for producing chimeric mammals that can be used to
XX CC develop new drugs and vaccine, factors, drugs and tissues for
XX CC transplantation, also useful to study human diseases. The present
XX CC sequence represents a nucleotide sequence given in the Sequence Listing
XX CC of the present invention but not mentioned further within the
XX CC specification.

<p> SQ Sequence 2294 BP; 466 A; 680 C; 641 G; 507 T; 0 U; 0 Other; Query Match 100.0%; Score 100; DB 12; Length 2294; Best Local Similarity 100.0%; Pred. No. 4.5e-26; Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0; Qy 1 GACGGATCGGGAGATCTCCCGATCCCGATCCCGATCGACTTCAGTACAATCTGCTCTGATG 60 Db 1 GACGGATCGGGAGATCTCCCGATCCCGATCCCGATCGACTTCAGTACAATCTGCTCTGATG 60 Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCGCTTGTGTGTT 100 Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCGCTTGTGTGTT 100 RESULT 6 AAV40006 ID AAV40006 standard; DNA; 3853 BP. XX AC AAV40006; XX DT 27-AUG-2003 (revised) DT 15-FEB-1999 (first entry) XX XX Plasmid pCTM. XX KW E2F; transcription factor; human; retinoblastoma protein RB; KW bladder cancer; restenosis; angioplasty; diabetic retinopathy; KW thyroid hyperplasia; Grave's disease; psoriasis; KW benign prostatic hypertrophy; Li-Fraumeni syndrome; KW peripheral vascular disease; therapy; plasmid pCTM; ss. XX OS Human cytomegalovirus. OS OS mastadenovirus. OS unidentified bacteriophage; T7. OS unidentified bacteriophage; SP6. OS Macaca mulatta; polyoma virus. OS Bos taurus. OS Chimeric. XX FH Key Location/Qualifiers FT promoter 209..864 FT /tag= a FT /note= "CMV promoter" FT misc_feature 907..1131 FT /tag= b FT /function= "tripartite leader sequence" FT promoter 1132..1149 FT /tag= c FT /note= "SP6 promoter" FT misc_feature 1679..3853 FT /tag= d FT /note= "pUC19 backbone H3 to AatII" FT CDS complement(2857..3717) FT /tag= e FT /note= "AMP-ORF" XX XX WO9821228-A1. XX XX 22-MAY-1998. XX XX 13-NOV-1997; 97WO-US021821. XX XX 15-NOV-1996; 96US-00751517. XX 14-FEB-1997; 97US-00801092. XX XX (CANJ-) CANJI INC. XX XX Antelman D, Gregory RJ, Wills KN; XX XX WPI; 1998-297858/26. XX XX New fusion polypeptide of, e.g. transcription factor - used to treat, </p>	<p> PT e.g. hyper-proliferative disease such as cancer and restenosis. XX XX Example 1; Fig 4; 91pp; English. XX CC This is the nucleotide sequence of pCTM, a plasmid which contains a CMV CC promoter, a tripartite adenovirus leader flanked by T7 and SP6 promoters, CC and a multiple cloning site with a bovine growth hormone polyA site and CC downstream SV40 polyA site. It has been used as a vector for the CC expression of fusion proteins of the invention that comprise CC retinoblastoma protein (BP, see AAW62465) and E2F transcription factor CC (see AAW62464). Such fusion proteins, particularly expressed from gene CC therapy vectors, are used to treat hyperproliferative conditions, they are CC specifically cancer (particularly of the bladder) or restenosis. They are CC more effective in repressing transcription of the E2F promoter than RB CC alone and cause cell-cycle arrest in a variety of cells. (Updated on 27- CC AUG-2003 to correct OS field.) XX SQ Sequence 3853 BP; 936 A; 986 C; 942 G; 989 T; 0 U; 0 Other; Query Match 100.0%; Score 100; DB 2; Length 3853; Best Local Similarity 100.0%; Pred. No. 5.2e-26; Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0 Qy 1 GACGGATCGGGAGATCTCCCGATCCCGATCCCGATCGACTTCAGTACAATCTGCTCTGATG 60 Db 1 GACGGATCGGGAGATCTCCCGATCCCGATCCCGATCGACTTCAGTACAATCTGCTCTGATG 60 Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCGCTTGTGTGTT 100 Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCGCTTGTGTGTT 100 RESULT 7 AAV40007 ID AAV40007 standard; DNA; 4026 BP. XX AC AAV40007; XX DT 27-AUG-2003 (revised) DT 15-FEB-1999 (first entry) XX XX Plasmid pCTMI. DE XX KW E2F; transcription factor; human; retinoblastoma protein RB; KW bladder cancer; restenosis; angioplasty; diabetic retinopathy; KW thyroid hyperplasia; Grave's disease; psoriasis; KW benign prostatic hypertrophy; Li-Fraumeni syndrome; KW peripheral vascular disease; therapy; plasmid pCTMI; ss. XX OS Human cytomegalovirus. OS OS mastadenovirus. OS unidentified bacteriophage; T7. OS unidentified bacteriophage; SP6. OS Macaca mulatta; polyoma virus. OS Bos taurus. OS Chimeric. XX FH Key Location/Qualifiers FT promoter 209..864 FT /tag= a FT /note= "CMV promoter" FT misc_feature 907..1074 FT /tag= b FT /function= "tripartite leader sequence" FT intron 1075..1253 FT /tag= c FT /note= "hybrid SV40 late intron" FT promoter 1305..1322 FT /tag= d FT /note= "SP6 promoter" FT misc_feature 1851..4026 FT /tag= e FT /note= "pUC19 backbone H3 to AatII" </p>
--	---

FT CDS complement(3032..3890)
 FT /*tag= f
 FT /*note= "AMP-ORF"
 PN W09821228-A1.
 XX PD 22-MAY-1998.
 XX PF 13-NOV-1997; 97WO-US021821.
 XX PR 15-NOV-1996; 96US-00751517.
 XX PR 14-FEB-1997; 97US-00801092.
 XX PA (CANJ-) CANJI INC.
 XX PI Antelman D, Gregory RJ, Wills KN;
 XX DR WPI; 1998-297858/26.
 XX PF New fusion polypeptide of, e.g. transcription factor - used to treat,
 FT e.g. hyper-proliferative disease such as cancer and restenosis.
 XX PS Example 1; Fig 6; 91pp; English.
 CC This is the nucleotide sequence of pCTMI, a plasmid that was constructed
 CC from pCTM (see AAV40006) by digesting pCTM with XhoI and NotI and
 CC subcloning a 180 bp intron XhoI-NotI fragment from a pCMV-beta-gal
 CC vector. Plasmid pCTMI has been used as a vector for the expression of
 CC fusion proteins of the invention that comprise retinoblastoma protein
 CC (BP, see AAW62465) and E2F transcription factor (see AAW62464). Such
 CC fusion proteins, particularly expressed from gene therapy vectors, are
 CC used to treat hyperproliferative conditions, specifically cancer
 CC (particularly of the bladder) or restenosis. They are more effective in
 CC repressing transcription of the E2F promoter than RB alone and cause cell
 CC -cycle arrest in a variety of cells. (Updated on 27-AUG-2003 to correct
 CC OS field.)
 XX SQ Query Match 4026 BP; 978 A; 1021 C; 982 G; 1045 T; 0 U; 0 Other;
 Best Local Similarity 100.0%; Score 100; DB 2; Length 4026;
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GACGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCATG 60
 Db 1 GACGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCATG 60
 Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTTT 100
 Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTTT 100
 RESULT 8
 AAV63466
 ID AAV63466 standard; DNA; 4249 BP.
 XX AC AAV63466;
 XX DT 27-AUG-2003 (revised)
 XX DT 15-FEB-1999 (first entry)
 XX DE Plasmid pCTMIE.
 XX E2F; transcription factor; human; retinoblastoma protein RB;
 KW bladder cancer; restenosis; angioplasty; diabetic retinopathy;
 KW thyroid hyperplasia; Grave's disease; psoriasis;
 KW benign prostatic hypertrophy; Li-Fraumeni syndrome;
 KW peripheral vascular disease; therapy; plasmid pCTMIE; ss.
 XX OS Human cytomegalovirus.
 OS mastadenovirus.
 OS unidentified bacteriophage; T7.
 OS unidentified bacteriophage; SP6.

OS Macaca mulatta; polyoma virus.
 OS Bos taurus.
 OS Chimeric.
 XX FH Key
 FT Promoter
 FT Location/Qualifiers
 FT 209..864
 FT /*tag= a
 FT /*note= "CMV promoter"
 FT 907..1074
 FT /*tag= b
 FT /*function= "tripartite leader sequence"
 FT 1081..1145
 FT /*tag= c
 FT /*note= "hybrid SV40 late intron"
 FT 1164..1366
 FT /*tag= d
 FT /*note= "early mRNA"
 FT 1261..1332
 FT /*tag= e
 FT /*note= "72 bp tandem repeat enhancer"
 FT 1333..1404
 FT /*tag= f
 FT /*note= "72 bp tandem repeat enhancer"
 FT 1366
 FT /*tag= g
 FT /*note= "T antigen binding site"
 FT 1372..1478
 FT /*tag= h
 FT /*note= "hybrid SV40 late intron"
 FT 1530..1545
 FT /*tag= i
 FT /*note= "SP6 promoter"
 FT 2075..4249
 FT /*tag= j
 FT /*note= "pUC19 backbone H3 to AatII"
 FT complement(3255..4113)
 FT /*tag= k
 FT /*note= "AMP-ORF"
 XX W09821228-A1.
 XX PD 22-MAY-1998.
 XX PF 13-NOV-1997; 97WO-US021821.
 XX PR 15-NOV-1996; 96US-00751517.
 XX PR 14-FEB-1997; 97US-00801092.
 XX PA (CANJ-) CANJI INC.
 XX PI Antelman D, Gregory RJ, Wills KN;
 XX DR WPI; 1998-297858/26.
 XX PF New fusion polypeptide of, e.g. transcription factor - used to treat,
 FT e.g. hyper-proliferative disease such as cancer and restenosis.
 XX PS Example 1; Fig 8; 91pp; English.
 CC This is the nucleotide sequence of pCTMIE, a plasmid that was constructed
 CC by amplifying the SV40 enhancer from SV40 viral DNA by PCR, digesting the
 CC amplified product with BglII and inserting into BamHI-digested plasmid
 CC pCTMI (see AAV40007). Plasmid pCTMIE has been used as a vector for the
 CC expression of fusion proteins of the invention that comprise
 CC retinoblastoma protein (BP, see AAW62465) and E2F transcription factor
 CC (see AAW62464). Such fusion proteins, particularly expressed from gene
 CC therapy vectors, are used to treat hyperproliferative conditions,
 CC specifically cancer (particularly of the bladder) or restenosis. They are
 CC more effective in repressing transcription of the E2F promoter than RB
 CC alone and cause cell-cycle arrest in a variety of cells. (Updated on 27-
 CC AUG-2003 to correct OS field.)
 XX SQ Sequence 4249 BP; 1020 A; 1074 C; 1048 G; 1107 T; 0 U; 0 Other;

```
Query Match      100.0%; Score 100; DB 2; Length 4249;
Best Local Similarity 100.0%; Pred. No. 5.4e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

[illegible]

RESULT 9
AAQ62391
ID AAQ62391 standard; DNA; 4341 BP.

AA
AC
AAQ62391;

	25-MAR-2003 (revised)	18-NOV-1994 (first entry)
DT		
DT		

Vector pVAC1.

Vector; pVAC1; pRc/RSV; leader sequence; termination signal;
KW fusion protein; pSfi/Not.Tag1; pS1B leader; human; immunoglobulin; VH1;
KW single chain; Fv; murine antibody; retroviral; envelope; plasmid;
KW vaccine; SS.

OS Synthetic.

	Key	Location/Qualifiers complement(1. .775)
FH		
FT	misc RNA	

FT /note= "Claim 9" /cag=

```
FT misc_mms
000: 0000
/*tag= b
```

FT misc RNA 606. .716

FT	/cag=	Claim 7"
FT	/note=	"Claim 7"

PN WO9408008-A1.

PD 14-APR-1994.

PF 04-OCT-1993; 93WO-GB002054.

PR 02-OCT-1992; 92GB-00020808.

PA (MEDI-) MEDICAL RES COUNCIL.

PI Hawkins RE, Russell SJ, Stevenson FK, Winter GP;

DR WPI; 1994-135575/16.

AA	
PT	Modulating immune response to a disease marker - by administering a
PT	vector which expresses the disease marker to interact with the immune
PT	system.

PS Claim 10: Fig 7: 77pp: English.

This sequence represents the vector pVAC1. This vector is based on the commercially available vector pRC/RSV. Leader sequences and termination signals were introduced into the vector to allow for production of fusion proteins. The vector, pSfi/Not-Tag1, was modified to replace the pCB1 leader with the human immunoglobulin VH1 leader sequence that permits the encoding of an SfiI cloning site without modification of the amino acid sequence. This fragment was then cloned as an EcoRI/Blunt-HindIII fragment into NotI/Blunt-HindIII cut vector pRC/RSV to give pVAC1. The single chain Fv for an individual patient can be inserted within the VH1 leader sequence. This plasmid when encoding a single chain murine

antibody/retroviral envelope fusion protein can be used as a plasmid vaccine and it induces a strong humoral response to the antibody moiety in BALB/c mice. (Updated on 25-MAR-2003 to correct PN field.)

Sequence 4341 BP; 1032 A; 1099 C; 1091 G; 1119 T; 0 U; 0 Other;

```
Query Match      100.0%; Score 100; DB 2; Length 4341;
Best Local Similarity 100.0%; Pred. No. 5.4e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 1 GACGGATCGGAGATCTCCCGATCCCGATCGACTCTCGACTCTCGATCAATCTGCTTGATG 60

Dδ 1 GACGGATCGGAGATCTCCCGATCCCGATCGACTCTCGACTCTCGATCAATCTGCTTGATG 60

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100

db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGT 100

RESULT 10

RESOL 10
AAS17704
ID AAS17704 standard; DNA; 4341 BP.

AA
AC AAS17704;

DT 12-MAR-2002 (first entry)

DE Vector pVAC1 encoding a DNA vaccine.

AA Cytostatic; vaccine; tetanus toxin; FrC; tumour; CTL; PCR primer; pVAc1;
 KW ds.

OS Clostridium tetani.

OS Homo sapiens.

OS Cauliflower mosaic virus.

PN WO200179510-A1.

PD 25-OCT-2001.

17-APR-2001; 2001WO-GB001719.

PR 17-APR-2000; 2000GB-00009470.

PA (CANC-) CANCER RES VENTURES LTD.

PI Rice J, Stevenson F;

DR WPI; 2002-066370/09.

Nucleic acid construct, useful to immunize against various diseases including cancer, expresses the first domain of tetanus toxin PrC fused to a disease peptide antigen to provide a vaccine.

PS Disclosure; Fig 4; 71pp; English.

The invention relates to a nucleic acid construct for delivery into living cells in vivo, to induce an immune response to a disease peptide antigen, where the construct directs expression of a fusion protein comprising the peptide antigen and the first domain of Fc. Also included are a nucleic acid vector comprising the above construct, a host cell comprising the above construct or vector and a method of producing a nucleic acid construct for inducing an immune response. The method comprises identifying a nucleic acid sequence encoding a disease peptide antigen comprising epitopes characteristic of the disease, cloning the nucleic acid sequence, introducing the cloned nucleic acid into a vector which allows the antigen to be expressed as a fusion with a first domain Fc from tetanus toxin, and optionally isolating the construct from the vector. The construct or vector is used as a vaccine to induce an immune response, particularly to tumour antigens. The present sequence is vector pVAC1 which encodes a vaccine of the invention

SQ Sequence 4341 BP; 1033 A; 1099 C; 1090 G; 1119 T; 0 U; 0 Other;
Query Match 100.0%; Score 100; DB 6; Length 4341;
Best Local Similarity 100.0%; Pred. No. 5.4e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GACGGATCGGAGATCTCCGATCCCTATGGTCGACTCTCAGTACAAATCTGCTCTGATG 60
DB |||||
XX |||||

OY 1 GACGGATCGGAGATCTCCGATCCCTATGGTCGACTCTCAGTACAAATCTGCTCTGATG 60
DB |||||
XX |||||

OY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
DB |||||
XX |||||

OY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
DB |||||
XX |||||

RESULT 11
ABN83143
ID ABN83143 standard; DNA; 4341 BP.
XX
AC ABN83143;
XX
DT 10-SEP-2002 (first entry)
XX
DE Plasmid pVAC1 complete sequence.
XX
KW Immune response; plant viral coat protein; pVAC1; cytostatic; virucide;
KW cancer; B cell malignancy; ds.
XX
OS Synthetic.
XX
FN WO200240513-A2.
XX
PD 23-MAY-2002.
XX
PF 20-NOV-2001; 2001WO-GB005142.
XX
PR 20-NOV-2000; 2000GB-00028319.
XX
PA (CANC-) CANCER RES VENTURES LTD.
XX
PI Savelyeva N, Stevenson F;
XX
XX WPI; 2002-500202/53.
XX
PS Nucleic acid construct for delivery into living cells as a vaccine,
PT useful for treating e.g. cancer, directs the expression of a fusion
PT protein comprising an antigen and an adjuvant sequence derived from a
PT plant viral coat protein.
XX
XX Example 3; Fig 7; 84pp; English.
XX
CC The invention relates to a novel nucleic acid construct for inducing an
CC immune response in vivo to an antigen, capable of directing the
CC expression of a fusion protein that comprises an antigen and an adjuvant
CC sequence derived from a plant viral coat protein. The construct of the
CC invention has cytostatic and virucide activity. The nucleic acid
CC construct is useful for inducing an immune response in a patient, for
CC vaccinating a patient against an infectious disease caused by an antigen
CC derived from a pathogen e.g. a virus, for treating a cancer patient or a
CC patient with a predisposition to cancer and for treating a patient having
CC a B cell malignancy, where the construct is encapsulated, and optionally,
CC a second nucleic acid sequence encoding a further immunomodulatory
CC polypeptide is administered to the patient. The construct is also useful
CC in medical treatment, and in the preparation of a vaccine for treating or
CC preventing a disease state associated with the antigen. The sequence
CC shows the complete sequence of vector pVAC1
XX
SQ Sequence 4341 BP; 1033 A; 1099 C; 1090 G; 1119 T; 0 U; 0 Other;
Query Match 100.0%; Score 100; DB 6; Length 4341;
Best Local Similarity 100.0%; Pred. No. 5.4e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GACGGATCGGAGATCTCCGATCCCTATGGTCGACTCTCAGTACAAATCTGCTCTGATG 60
DB |||||
XX |||||

OY 1 GACGGATCGGAGATCTCCGATCCCTATGGTCGACTCTCAGTACAAATCTGCTCTGATG 60
DB |||||
XX |||||

OY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
DB |||||
XX |||||

OY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
DB |||||
XX |||||

RESULT 12
AAF24901
ID AAF24901 standard; DNA; 4597 BP.
XX
AC AAF24901;
XX
DT 20-APR-2001 (first entry)
XX
DE Nucleotide sequence of the plasmid pCDNA3.1/GS.
XX
KW Microsphere; dihydrazide; hyaluronic acid; inflammatory response;
KW myocardial ischemia; cardiac angiogenesis; haemophilia;
KW vascular endothelial growth factor; VEGF; ss.
XX
OS Synthetic.
XX
FN WO200078358-A2.
XX
PD 28-DEC-2000.
XX
PF 19-JUN-2000; 2000WO-US016837.
XX
PR 18-JUN-1999; 99US-0140260P.
XX
PA (COLL-) COLLABORATIVE GROUP LTD.
XX
PI Chen W;
XX
XX WPI; 2001-071363/08.
XX
DR Hyaluronic acid micro spheres for use in gene therapy of myocardial
PT ischemia and hemophilia, comprising dihydrazide derivatized hyaluronic
PT acids crosslinked to nucleic acids.
XX
XX Example 1; Page 36-38; 38pp; English.
XX
CC The specification describes a microsphere comprising dihydrazide
CC derivatized hyaluronic acid crosslinked to a nucleic acid (NA). The
CC microspheres cause reduced inflammatory responses, and have increased
CC safety and biodegradability. The microspheres are useful for transfecting
CC a cell of a subject and for treating a subject having myocardial
CC ischemia, by increasing cardiac angiogenesis. They are also useful for
CC treating haemophilia. The present sequence represents the plasmid
CC pCDNA3.1/GS, into which is inserted a polynucleotide sequence which is
CC crosslinked to hyaluronic acid. The polynucleotide sequence encodes a
CC vascular endothelial growth factor (VEGF)
XX
SQ Sequence 4597 BP; 1062 A; 1214 C; 1206 G; 1115 T; 0 U; 0 Other;
Query Match 100.0%; Score 100; DB 4; Length 4597;
Best Local Similarity 100.0%; Pred. No. 5.5e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GACGGATCGGAGATCTCCGATCCCTATGGTCGACTCTCAGTACAAATCTGCTCTGATG 60
DB |||||
XX |||||

OY 1 GACGGATCGGAGATCTCCGATCCCTATGGTCGACTCTCAGTACAAATCTGCTCTGATG 60
DB |||||
XX |||||

OY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
DB |||||
XX |||||

OY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
DB |||||
XX |||||

RESULT 13
AAD39652

```

ID AD39652 standard; DNA; 4639 BP.
XX
AC AAD39652;
XX
XX 22-OCT-2002 (first entry)
DT
DE Human small nuclear RNA (snRNA) DNA.
XX
XX Human; recombinant vector; insertion cassette; small nuclear RNA; snRNA;
KW transgenic animal; ds.
XX
XX Homo sapiens.
OS
XX US2002058287-A1.
PN
XX 16-MAY-2002.
PD
XX
XX 12-MAR-2001; 2001US-00804481.
PF
XX
XX 10-MAR-2000; 2000US-0188304P.
PR
XX (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX
XX Graaf DD, Lander ES;
PI
XX WPI; 2002-499510/53.
XX
XX New recombinant vector containing sequence for small nuclear RNA, useful
PT e.g. for identifying variant snRNA that suppresses expression of
PT transcription products.
XX
XX Disclosure; Fig 1; 18pp; English.
XX
XX The invention relates to a recombinant vector which comprises DNA,
CC consisting of an insertion cassette contained between at least two
CC insertion sites, that encodes a small nuclear (sn) RNA. The invention is
CC used to identify snRNA modifications that inhibit expression of
CC transcription products (and the identified snRNA are used to suppress
CC expression) for delivering antisense sequences to the nucleus and to
CC create transgenic animals. The present DNA sequence is human snRNA, U1
XX
XX Sequence 4639 BP; 1067 A; 1198 C; 1243 G; 1131 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 100; DB 6; Length 4639;
Best Local Similarity 100.0%; Pred. No. 5.5e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GACGATCGGAGATCTCCGATCCCTATGCTCGACTTCTCAGTACAAATCTGCTCTGATG 60
Db 1 GACGATCGGAGATCTCCGATCCCTATGCTCGACTTCTCAGTACAAATCTGCTCTGATG 60
Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
RESULT 14
AAF83146
ID AAF83146 standard; DNA; 4840 BP.
XX
AC AAF83146;
XX
XX 09-JUL-2001 (first entry)
DT
DE Complete sequence of vector pIRES/BS.
XX
XX Blastocidin resistance; BS gene; gene therapy; tissue engineering;
KW cosmetic surgery; arthritis; joint replacement; skin graft; rhinoplasty;
KW pIRES/BS; ss.
XX
XX Synthetic.
OS
XX WO200119853-A2.
PN
AD39652 standard; DNA; 4639 BP.
XX
AC AAD39652;
XX
XX 22-OCT-2002 (first entry)
DT
DE Human small nuclear RNA (snRNA) DNA.
XX
XX Human; recombinant vector; insertion cassette; small nuclear RNA; snRNA;
KW transgenic animal; ds.
XX
XX Homo sapiens.
OS
XX US2002058287-A1.
PN
XX 16-MAY-2002.
PD
XX
XX 12-MAR-2001; 2001US-00804481.
PF
XX
XX 10-MAR-2000; 2000US-0188304P.
PR
XX (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX
XX Graaf DD, Lander ES;
PI
XX WPI; 2002-499510/53.
XX
XX New recombinant vector containing sequence for small nuclear RNA, useful
PT e.g. for identifying variant snRNA that suppresses expression of
PT transcription products.
XX
XX Disclosure; Fig 1; 18pp; English.
XX
XX The invention relates to a recombinant vector which comprises DNA,
CC consisting of an insertion cassette contained between at least two
CC insertion sites, that encodes a small nuclear (sn) RNA. The invention is
CC used to identify snRNA modifications that inhibit expression of
CC transcription products (and the identified snRNA are used to suppress
CC expression) for delivering antisense sequences to the nucleus and to
CC create transgenic animals. The present DNA sequence is human snRNA, U1
XX
XX Sequence 4639 BP; 1067 A; 1198 C; 1243 G; 1131 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 100; DB 6; Length 4639;
Best Local Similarity 100.0%; Pred. No. 5.5e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GACGATCGGAGATCTCCGATCCCTATGCTCGACTTCTCAGTACAAATCTGCTCTGATG 60
Db 1 GACGATCGGAGATCTCCGATCCCTATGCTCGACTTCTCAGTACAAATCTGCTCTGATG 60
Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
RESULT 15
ADB33528
ID ADB33528 standard; DNA; 5015 BP.
XX
AC ADB33528;
XX
XX 04-DEC-2003 (first entry)
DT
DE Expression vector nucleotide sequence SEQ ID NO:27.
XX
XX fusion protein; amyloid precursor protein; APP; transcription factor;
KW neurotrophic; neuroprotective; APP inhibitor;
KW amyloid precursor protein inhibitor; Alzheimer's disease; beta-secretase;
KW gamma-secretase; human; gene; ds.
XX
XX Synthetic.
OS
XX Homo sapiens.
XX
XX WO2003072041-A2.
PN
XX 04-SEP-2003.
XX
XX 23-FEB-2003; 2003WO-US005458.
PF

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 05:15:57 ; Search time 952.146 Seconds
(without alignments)
3997.736 Million cell updates/sec

Title: US-09-482-682-47_COPY_1_100

Perfect score: 100

Sequence: 1 gacggatcggagatctccc.....ctgtccctgtgtgtgtt 100

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	60.0	602	8	B67169 CpG0047A Cp
2	55.6	55.6	694	8	B2052929 jnr13g03.
3	55.6	55.6	696	8	B2050328 jnr42c12.
4	55.6	55.6	717	8	BZ054067 jnr38b09.
5	53.6	53.6	348	2	AW409112 sal10h5 S
6	53.4	53.4	343	1	AL715724 AL715724
7	53.4	53.4	345	1	AL714571 AL714571
8	53.4	53.4	761	7	CK119397 212o09.pl
9	53.4	53.4	766	7	CK120360 207j04.pl
10	53.4	53.4	788	7	CK117844 209p08.pl
11	53.4	53.4	898	9	CL141237 ISB1-118J
12	53.4	53.4	899	9	CL140877 ISB1-118B
13	53.4	53.4	1009	9	CL123953 ISB1-84J1
14	53.2	53.2	814	8	AQ914559 nbe0049M
15	53	53.0	675	8	BZ051815 jnr57D03.
16	53	53.0	679	8	BZ052857 jnr13g03.
17	53	53.0	700	8	BZ050646 jnr66f08.
18	53	53.0	701	8	BZ052015 jnr56b03.
19	53	53.0	708	8	BZ054793 jnr33g03.
20	53	53.0	709	8	BZ053587 jnr98D01.
21	53	53.0	712	8	BZ054005 jnr38b09.
22	52.8	52.8	451	8	AQ863966 nbe0002E
23	52.6	52.6	399	8	AQ075099 CIT-HSP-2
24	52.4	52.4	700	8	BZ049113 jnr21d02.

25	52.4	52.4	708	8	BZ050047
26	51.6	51.6	328	9	CC819886
27	51.6	51.6	351	9	CC818492
28	51.6	51.6	358	9	CC817661
29	51.6	51.6	364	9	CC817805
30	51.6	51.6	364	9	CC818511
31	51.6	51.6	364	9	CC818574
32	51.6	51.6	364	9	CC819049
33	51.6	51.6	369	9	CC817069
34	51.6	51.6	374	9	CC817074
35	51.6	51.6	374	9	CC820036
36	51.6	51.6	395	9	CC817652
37	51.6	51.6	403	9	CC817682
38	51.6	51.6	403	9	CC817837
39	51.6	51.6	414	9	CC819240
40	51.6	51.6	419	9	CC818384
41	51.6	51.6	420	9	CC817834
42	51.6	51.6	426	9	CC817720
43	51.6	51.6	437	9	CC819820
44	51.6	51.6	441	9	CC818421
45	51.6	51.6	443	9	CC817769

ALIGNMENTS

RESULT 1
LOCUS B67169
DEFINITION CpG0047A CpIOWAgDNA2 Cryptosporidium parvum genomic, GSS 12-MAY-2000
sequence.
ACCESSION B67169
VERSION B67169.1 GI:2642750
KEYWORDS GSS.
SOURCE Cryptosporidium parvum
ORGANISM Cryptosporidium parvum
REFERENCE 1 (bases 1 to 602)
AUTHORS Strong, W.B. and Nelson, R.G.
TITLE Preliminary profile of the Cryptosporidium parvum genome: an expressed sequence tag and genome survey sequence analysis
JOURNAL Mol. Biochem. Parasitol. 107 (1); 1-32 (2000)
MEDLINE 20183851
PUBMED 10717299
COMMENT Contact: Nelson, R. G.
Depts. of Medicine & Pharmaceutical Chemistry
San Francisco General Hospital-University of California, San Francisco
Box 0811, San Francisco, CA 94143-0811, USA
Tel: 415 206 8946
Fax: 415 206 3353
Email: malariaditsa.ucsf.edu
Submitted sequence has been edited to remove vector sequences 5' to the insert, to correct miscalled bases and assign uncalled (N) bases throughout the sequence, and to terminate when base-calling became ambiguous.
Seq primer: T7
Class: shotgun
High quality sequence stop: 602.

FEATURES
source
1..602
Location/Qualifiers
/organism="Cryptosporidium parvum"
/mol_type="genomic DNA"
/strain="IOWA"
/db_xref="taxon:5807"
/lab_host="E. coli XL2 Blue MRF"
/clone_lib="CpIOWAgDNA2"
/note="vector: PCR-script Amp SK+; Site_1: SrfI; C. parvum (IOWA isolate) genomic DNA was hydrodynamically sheared to produce fragments having a tight size distribution between 2-4 kb by Dr. Yvonne Thorstenson of the Stanford DNA Sequencing and Technology Center

(http://sequence-www.stanford.edu/group/techdev/shear.htm)

The randomly sheared gDNA was chromatographed on Sephacryl S-400 to remove any small fragments and DNA eluting in the void volume was blunt-ended with T4 DNA Polymerase, treated with alkaline phosphatase to prevent the ligation of multiple fragments, ligated to Srf I-digested pCR-Script Amp (SK+) vector and transformed into E. coli strain XL10 Gold. Recombinant clones from the first plating of the library were selected for sequence analysis using T3 and T7 primers."

ORIGIN

Query Match 60.0%; Score 60; DB 8; Length 602;
 Best Local Similarity 100.0%; Pred. No. 2.4e-10;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 41 CAGTACAACTGCTCTGATGCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100
 |||||
 Db 1 CAGTACAACTGCTCTGATGCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 60
 |||||

RESULT 2
 BZ052929/c
 LOCUS jnr13903.g1 B.oleracea001 Brassica oleracea genomic, genomic survey
 DEFINITION sequence.

ACCESSION BZ052929
 VERSION BZ052929.1 GI:23654922
 KEYWORDS GSS.
 SOURCE Brassica oleracea
 ORGANISM Brassica oleracea

REFERENCE
 AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

1 (bases 1 to 694)
 Delehaunty, K., Fewell, G., Fulton, L., McCombie, W.R., Miner, T., Nash, W., Rabinowicz, P.D. and Wilson, R.K.
 Whole genome shotgun reads from Brassica oleracea
 Unpublished (2002)
 CONTACT: Richard K. Wilson
 Genome Sequencing Center
 Washington University School of Medicine
 Email: submissions@watson.wustl.edu
 Plate: jnr13 row: 9 column: 03
 Seq primer: -28RppOT reverse
 Class: shotgun
 High quality sequence start: 32
 High quality sequence stop: 551.

FEATURES
 source Location/Qualifiers
 1..694
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /db_xref="taxon:3712"
 /clone_lib="B.oleracea001"
 /note="Vector: pOTw13; Whole genome shotgun library from flowering buds. DNA was purified from a crude nuclear prep using Brassica oleracea T01000DH3 buds provided by Thomas Osborn at the University of Wisconsin. Genomic DNA was provided by Pablo Rabinowicz (CSHL) and the shotgun library prepared at Washington University Genome Sequencing Center."

Query Match 55.6%; Score 55.6; DB 8; Length 694;
 Best Local Similarity 77.9%; Pred. No. 9e-09;
 Matches 67; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 3 CGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATGCC 62
 |||||
 Db 324 CGGATCGATAGGTCCTCGGACTAGTTATGTCGACTCTCAGTACAATCTGCTCTGATGCC 265
 |||||

Qy 63 GCATAGTTAAGCCAGTATCTGCTCCC 88
 |||||

ORIGIN

Query Match 55.6%; Score 55.6; DB 8; Length 694;
 Best Local Similarity 77.9%; Pred. No. 9e-09;
 Matches 67; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 3 CGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATGCC 62
 |||||
 Db 324 CGGATCGATAGGTCCTCGGACTAGTTATGTCGACTCTCAGTACAATCTGCTCTGATGCC 265
 |||||

Qy 63 GCATAGTTAAGCCAGTATCTGCTCCC 88
 |||||

Db 264 GCATAGTTAAGCCAGCCCGACACCC 239

RESULT 3

BZ050328

LOCUS

DEFINITION

sequence.

ACCESSION BZ050328

VERSION BZ050328.1 GI:23649718

KEYWORDS GSS.

SOURCE Brassica oleracea

ORGANISM Brassica oleracea

REFERENCE

AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

1 (bases 1 to 696)

Delehaunty, K., Fewell, G., Fulton, L., McCombie, W.R., Miner, T., Nash, W., Rabinowicz, P.D. and Wilson, R.K.

Whole genome shotgun reads from Brassica oleracea

Unpublished (2002)

CONTACT: Richard K. Wilson

Genome Sequencing Center

Washington University School of Medicine

Email: submissions@watson.wustl.edu

Plate: jnr42 row: C column: 12

Seq primer: -21UPPOT forward

Class: shotgun

High quality sequence start: 35

High quality sequence stop: 180.

FEATURES

source Location/Qualifiers

1..696

/organism="Brassica oleracea"

/mol_type="genomic DNA"

/db_xref="taxon:3712"

/clone_lib="B.oleracea001"

/note="Vector: pOTw13; Whole genome shotgun library from flowering buds. DNA was purified from a crude nuclear prep using Brassica oleracea T01000DH3 buds provided by Thomas Osborn at the University of Wisconsin. Genomic DNA was provided by Pablo Rabinowicz (CSHL) and the shotgun library prepared at Washington University Genome Sequencing Center."

ORIGIN

Query Match 55.6%; Score 55.6; DB 8; Length 696;

Best Local Similarity 77.9%; Pred. No. 9e-09;

Matches 67; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 3 CGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATGCC 62

|||||

Db 45 CGGATCGATAGGTCCTCGGACTAGTTATGTCGACTCTCAGTACAATCTGCTCTGATGCC 104

|||||

Qy 63 GCATAGTTAAGCCAGTATCTGCTCCC 88

|||||

Db 105 GCATAGTTAAGCCAGCCCGACACCC 130

|||||

RESULT 4

BZ054067/c

LOCUS

DEFINITION

sequence.

ACCESSION BZ054067

VERSION BZ054067.1 GI:23657216

KEYWORDS GSS.

SOURCE Brassica oleracea

ORGANISM Brassica oleracea

REFERENCE

AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

1 (bases 1 to 717)

Delehaunty, K., Fewell, G., Fulton, L., McCombie, W.R., Miner, T.,

Whole genome shotgun reads from Brassica oleracea

Unpublished (2002)

CONTACT: Richard K. Wilson

Genome Sequencing Center

Washington University School of Medicine

Email: submissions@watson.wustl.edu

Plate: jnr38 row: 9 column: 03

Seq primer: -28RppOT reverse

Class: shotgun

High quality sequence start: 32

High quality sequence stop: 551.

FEATURES

source Location/Qualifiers

1..694

/organism="Brassica oleracea"

/mol_type="genomic DNA"

/db_xref="taxon:3712"

/clone_lib="B.oleracea001"

/note="Vector: pOTw13; Whole genome shotgun library from flowering buds. DNA was purified from a crude nuclear prep using Brassica oleracea T01000DH3 buds provided by Thomas Osborn at the University of Wisconsin. Genomic DNA was provided by Pablo Rabinowicz (CSHL) and the shotgun library prepared at Washington University Genome Sequencing Center."

TITLE Nash, W., Rabinowicz, P.D. and Wilson, R.K.
JOURNAL Whole genome shotgun reads from Brassica oleracea
COMMENT Unpublished (2002)
Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Plate: jnr38 row: b column: 09
Seq primer: -28RPpOT reverse
Class: shotgun
High quality sequence start: 87
High quality sequence stop: 543.
Location/Qualifiers

FEATURES

source

1. .717
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/db_xref="taxon:3712"
/clone_lib="B.oleracea001"
/note="Vector: pOTw13; Whole genome shotgun library from
flowering buds. DNA was purified from a crude nuclear
prep using Brassica oleracea T01000DH3 buds provided by
Thomas Osborn at the University of Wisconsin. Genomic
DNA was provided by Pablo Rabinowicz (CSHL) and the
shotgun library prepared at Washington University Genome
Sequencing Center."

ORIGIN

Query Match 55.6%; Score 55.6; DB 8; Length 717;
Best Local Similarity 77.9%; Pred. No. 9.1e-09;
Matches 67; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 3 CGGATCGGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTGTGATGCC 62

Db 248 CGGATCGATAGTCCCTCGACTAGTATGTCGACTCTCAGTACAAATCTGCTGTGATGCC 189

Qy 63 GCATAGTTAAGCAGATATCTGCTCC 88

Db 188 GCATAGTTAAGCAGCCCGACACC 163

RESULT 5

AW409112
LOCUS sal10h5 Salivary Gland Library Homo sapiens linear EST 31-DEC-2000
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 348)

Bergheim, A., Ogilvie, E., Arndt, S., Napier, H., Taylor, M., Lovett, M.,
Simmons, A., Hide, W. and Ramsay, M.

A high density transcript map between markers D8S550 and D8S1759 on
8p22-23, using cDNA direct selection
Unpublished (2000)

Contact: Ramsey M

Department of Human Genetics
South African Institute For Medical Research
P.O.Box 1038, Johannesburg, Gauteng, 2000, South Africa
Fax: 2711 489 9226
Email: micheler@mail.saimr.wits.ac.za.

FEATURES

source

1. .348
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/issue_type="Salivary Gland"
/clone_lib="Salivary Gland Library"
/note="Vector: pAMP10"

ORIGIN

Query Match 53.6%; Score 53.6; DB 2; Length 348;
Best Local Similarity 80.5%; Pred. No. 4.1e-08;
Matches 62; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 11 GAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTGTGATCCGCGATAGTT 70

Db 65 GCGGTATACACCGCATATGGTGCCTCTCAGTACAAATCTGCTGTGATCCGCGATAGTT 124

Qy 71 AAGCCAGTATCTGCTCC 87

Db 125 AAGCCAGTATACACTCC 141

RESULT 6

AL715724/c

LOCUS

DEFINITION AL715724 Danio rerio embryonic inner ear subtracted cDNA Danio
rerio cDNA clone BN0AA018ZF12 5', mRNA sequence.

ACCESSION AL715724

VERSION AL715724.1 GI:20180327

KEYWORDS EST.

SOURCE Danio rerio (zebrafish)

ORGANISM

Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Ostariophysi;

Cypriniformes; Cyprinidae; Danio.

1 (bases 1 to 343)

Coimbra, R., Weil, D., Brottier, P., Blanchard, S., Levi, M.,

Hardelin, J.P., Weissenbach, J. and Petit, C.

A subtracted cDNA library from the zebrafish (Danio rerio)

embryonic inner ear

Unpublished (2002)

Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

source

1. .343
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="BN0AA018ZF12"
/tissue_type="inner ear"
/dev_stage="embryonic"
/clone_lib="Danio rerio embryonic inner ear subtracted
cDNA"
/note="subtracted cDNA library"

ORIGIN

Query Match 53.4%; Score 53.4; DB 1; Length 343;
Best Local Similarity 84.5%; Pred. No. 4.8e-08;
Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 17 TCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTGTGATCCGCGATAGTTAAGCCA 76

Db 280 TTCACACCGCATATGGTGCCTCTCAGTACAAATCTGCTGTGATCCGCGATAGTTAAGCCA 221

Qy 77 GTATCTGCTCC 87

Db 220 GTATACACTCC 210

RESULT 7

AL714571/c

LOCUS

DEFINITION AL714571 Danio rerio embryonic inner ear subtracted cDNA Danio
rerio cDNA clone BN0AA007ZC02 5', mRNA sequence.

ACCESSION AL714571

VERSION AL714571.1 GI:20179174

KEYWORDS EST.

SOURCE Danio rerio (zebrafish)

ORGANISM

Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE
AUTHORS 1 (bases 1 to 345)
          Coimbra, R., Weil, D., Brottier, P., Blanchard, S., Levi, M.,
          Hardelin, J.P., Weissenbach, J. and Petit, C.
TITLE    A subtracted cDNA library from the zebrafish (Danio rerio)
          embryonic inner ear
JOURNAL  Embryonic inner ear
COMMENT  Unpublished (2002)
CONTACT: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr.
FEATURES
source   Location/Qualifiers
1..345   /organism="Danio rerio"
         /mol_type="mRNA"
         /db_xref="taxon:7955"
         /clone="BN0AA0072C02"
         /tissue_type="inner ear"
         /dev_stage="embryonic"
         /clone_lib="Danio rerio embryonic inner ear subtracted
         cDNA"
         /note="subtracted cDNA library"

ORIGIN
Query Match      53.4%; Score 53.4; DB 1; Length 345;
Best Local Similarity 84.5%; Pred. No. 4.8e-08;
Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 17 TCCCGATCCCTATGTCGACTCTCAGTACAACTGCTCTGATGCCGCGCATAGTTAAGCCA 76
Db 280 TTCACCGCATATGGTGCACTCTCAGTACAACTGCTCTGATGCCGCGCATAGTTAAGCCA 221

Qy 77 GTATCTGCTCC 87
Db 220 GTATACACTCC 210

RESULT 8
CK119397/c
LOCUS    CK119397
DEFINITION 212009, pl AtM1 Arabidopsis thaliana cDNA clone MPMGp2011J009212
5-PRIME, mRNA sequence.
ACCESSION CK119397
VERSION    CK119397.1 GI:47829713
KEYWORDS   EST.
SOURCE     Arabidopsis thaliana (thale cress)
ORGANISM   Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
AUTHORS    Feilner, T., Immink, R.G.H., Cahill, D.J. and Kersten, B.
TITLE      Generation of a cDNA expression library from Arabidopsis
inflowence meristem
JOURNAL    Unpublished (2003)
COMMENT    Contact: Birgit Kersten
           Plant Protein Chip Group, Department Lehrach
           Max-Planck-Institute for Molecular Genetics
           Innestr. 73 D-14195 Berlin, Germany
           Tel: +49(0)30/84131648
           Fax: +49(0)30/84131128
           Email: Kersten@molgen.mpg.de
           Insert Length: 761 Std Error: 0.00
           Plate: 212 row: 0 column: 9
           Seq primer: POE65
           Location/Qualifiers
           1..761
           /organism="Arabidopsis thaliana"
           /mol_type="mRNA"
           /ecotype="Columbia"
           /db_xref="GABI:954234"
           /db_xref="taxon:3702"

Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE
AUTHORS 1 (bases 1 to 345)
          Coimbra, R., Weil, D., Brottier, P., Blanchard, S., Levi, M.,
          Hardelin, J.P., Weissenbach, J. and Petit, C.
TITLE    A subtracted cDNA library from the zebrafish (Danio rerio)
          embryonic inner ear
JOURNAL  Embryonic inner ear
COMMENT  Unpublished (2002)
CONTACT: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr.
FEATURES
source   Location/Qualifiers
1..345   /organism="Danio rerio"
         /mol_type="mRNA"
         /db_xref="taxon:7955"
         /clone="BN0AA0072C02"
         /tissue_type="inner ear"
         /dev_stage="embryonic"
         /clone_lib="Danio rerio embryonic inner ear subtracted
         cDNA"
         /note="subtracted cDNA library"

ORIGIN
Query Match      53.4%; Score 53.4; DB 1; Length 345;
Best Local Similarity 84.5%; Pred. No. 4.8e-08;
Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 17 TCCCGATCCCTATGTCGACTCTCAGTACAACTGCTCTGATGCCGCGCATAGTTAAGCCA 76
Db 280 TTCACCGCATATGGTGCACTCTCAGTACAACTGCTCTGATGCCGCGCATAGTTAAGCCA 221

Qy 77 GTATCTGCTCC 87
Db 220 GTATACACTCC 210

RESULT 9
CK120360/c
LOCUS    CK120360
DEFINITION 207504, pl AtM1 Arabidopsis thaliana cDNA clone MPMGp2011J04207
5-PRIME, mRNA sequence.
ACCESSION CK120360
VERSION    CK120360.1 GI:47830676
KEYWORDS   EST.
SOURCE     Arabidopsis thaliana (thale cress)
ORGANISM   Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
AUTHORS    Feilner, T., Immink, R.G.H., Cahill, D.J. and Kersten, B.
TITLE      Generation of a cDNA expression library from Arabidopsis
inflowence meristem
JOURNAL    Unpublished (2003)
COMMENT    Contact: Birgit Kersten
           Plant Protein Chip Group, Department Lehrach
           Max-Planck-Institute for Molecular Genetics
           Innestr. 73 D-14195 Berlin, Germany
           Tel: +49(0)30/84131648
           Fax: +49(0)30/84131128
           Email: Kersten@molgen.mpg.de
           Insert Length: 766 Std Error: 0.00
           Plate: 207 row: J column: 4
           Seq primer: POE65
           Location/Qualifiers
           1..766
           /organism="Arabidopsis thaliana"
           /mol_type="mRNA"
           /ecotype="Columbia"
           /db_xref="GABI:953059"
           /db_xref="taxon:3702"
           /clone="MPMGp2011J04207"
           /tissue_type="inflowence meristem"
           /dev_stage="about one week after bolting"
           /lab_host="E. coli SCS-1/pSE111"
           /clone_lib="AtM1"
           /note="Vector: POE-30NAST-attB (AY386205); Site_1: SalI;

```

expression clones. Average insert size is 1 kb. Note: The rearranged sublibrary (plate numbers begin with 201) was sequenced. Library generation and sequencing was granted in context of GABI-LAPP; data are also accessible at <https://gabi.rzpd.de>

17 TCCGATCCCTATGGTGAATCTCTCAGTACAAATCTGCTCTGATGCCGATAGTTAAGCCA 76
 Qy |||||
 514 TTACACCGCATATGGTGACTCTCAGTACAAATCTGCTCTGATGCCGATAGTTAAGCCA 45
 Db |||||
 77 GTATCTGTCC 87
 Qy |||||
 454 GTATACATCC 444
 Db |||||

RESULT 11
CL141237/0

RESULT 11					
CL141237/c					
LOCUS	CL141237	898 bp	DNA	linear	GSS 05-JAN-2004
DEFINITION	ISB1-118J17	T7.1 ISB1	Xenopus tropicalis	genomic clone	ISB1-118J17,
				genomic survey sequence.	

ACCESSION CL141237
VERSION CL141237.1
KEYWORDS GSS.

ACCESSION	CL141237	GI:40634872
VERSION	CL141237.1	
KEYWORDS	GSS.	
SOURCE	Xenopus tropicalis (western clawed frog)	
ORGANISM	Xenopus tropicalis	
	Xenopus tropicalis	
	Eukaryota; Metazoa; Chordata; Craniata;	
	Amphibia; Batrachia; Anura; Mesobatrachii	
	Xenopodinae; Xenopus; Silurana.	
REFERENCE	1 (bases 1 to 898)	
AUTHORS	Krenitski,C., Carter,J., McPherson,J., W Mardis,E. and Willson,R.	
TITLE	A physical map of the xenopus tropicalis	
JOURNAL	Unpublished (2003)	
COMMENT	Contact: Richard K Wilson	
	Genome Sequencing Center	
	Washington University School of Medicine	
	Email: submissions@watson.wustl.edu	
	Insert Length: 75000 Std Error: 0.00	
	Seq primer: 17 TAATACGACTCACTATAGG	
	Class: BAC ends	
	High quality sequence start: 4	
	High quality sequence stop: 742.	

FEATURES
SOURCE

```

FEATURES
    source
        1..898
            /organism="Xenopus tropicalis"
            /mol_type="genomic DNA"
            /db_xref="taxon:8364"
            /clone="ISB1-118J17"
            /clone_lib="ISB1"
            /note="Vector: pBelOBAcl1; ISB-1 Xenopus tropicalis BAC
            Library Segment 1"

```

Qy	17	TCCGATCCCTATGGTGCAGTCTCAGTACAAATCTGCTCTGATGCCGCGATAGTTAAGCCA	76
Db	195	TTCAACCGCATATGGTGCAGTCTCAGTACAAATCTGCTCTGATGCCGCGATAGTTAAGCCA	13
Qy	77	GTATCTGTCTCC	87
Db	135	GTATACACTCC	125

```

LOCUS      CL140877              899 bp    DNA          linear    GSS 05-JAN-2004
DEFINITION ISB1-118B12_T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-118B12,
            genomic survey sequence.
ACCESSION  CL140877
VERSION    CL140877.1  GI:40634512
KEYWORDS   GSS.
SOURCE     Xenopus tropicalis (western clawed frog)
ORGANISM   Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE  1 (bases 1 to 899)
            Kremitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
            Mardis, E. and Wilson, R.
            A physical map of the xenopus tropicalis genome
            Unpublished (2003)
            Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 75000 Std Error: 0.00
            Seq primer: T7 TAATACGACTCACTATAGG
            Class: BAC ends
            High quality sequence start: 4
            High quality sequence stop: 681.
FEATURES   source
            Location/Qualifiers
                1..899
                /organism="Xenopus tropicalis"
                /mol_type="genomic DNA"
                /db_xref="taxon:8364"
                /clone="ISB1-118B12"
                /clone_lib="ISB1"
                /notes="Vector: pBeloBAC11; ISB-1 Xenopus tropicalis BAC
                Library Segment 1"
ORIGIN
            Query Match      53.4%; Score 53.4; DB 9; Length 899;
            Best Local Similarity 84.5%; Pred. No. 5.8e-08;
            Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
Qy      17  TCCGATCCCTATGGTGCATCTCTCAGTACAACTGCTCTGATGCCGATAGTTAAGCCA 76
            | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      195  TTCACCGCATATGGTGCATCTCTCAGTACAACTGCTCTGATGCCGATAGTTAAGCCA 136
            | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy      77  GTATCTGCTCC 87
            | | | | | | | |
Db      135  GTATACACTCC 125

RESULT 13
LOCUS      CL123953/c              1009 bp    DNA          linear    GSS 05-JAN-2004
DEFINITION ISB1-94J15_T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-84J15,
            genomic survey sequence.
ACCESSION  CL123953
VERSION    CL123953.1  GI:40617588
KEYWORDS   GSS.
SOURCE     Xenopus tropicalis (western clawed frog)
ORGANISM   Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE  1 (bases 1 to 1009)
            Kremitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
            Mardis, E. and Wilson, R.
            A physical map of the xenopus tropicalis genome
            Unpublished (2003)
            Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 75000 Std Error: 0.00
            Seq primer: T7 TAATACGACTCACTATAGG

```

```

Class: BAC ends
High quality sequence start: 167
High quality sequence stop: 324.
FEATURES   source
            Location/Qualifiers
                1..1009
                /organism="Xenopus tropicalis"
                /mol_type="genomic DNA"
                /db_xref="taxon:8364"
                /clone="ISB1-84J15"
                /clone_lib="ISB1"
                /notes="Vector: pBeloBAC11; ISB-1 Xenopus tropicalis BAC
                Library Segment 1"
ORIGIN
            Query Match      53.4%; Score 53.4; DB 9; Length 1009;
            Best Local Similarity 84.5%; Pred. No. 5.9e-08;
            Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
Qy      17  TCCGATCCCTATGGTGCATCTCTCAGTACAACTGCTCTGATGCCGATAGTTAAGCCA 76
            | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      252  TTCACCGCATATGGTGCATCTCTCAGTACAACTGCTCTGATGCCGATAGTTAAGCCA 193
            | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy      77  GTATCTGCTCC 87
            | | | | | | | |
Db      192  GTATACACTCC 182

RESULT 14
LOCUS      AQ914559              814 bp    DNA          linear    GSS 02-DEC-1999
DEFINITION nhe0049W21r CUGI Rice BAC Library (ECORI) Oryza sativa (japonica
            cultivar-group) genomic clone nhe0049W21r, genomic survey
            sequence.
ACCESSION  AQ914559
VERSION    AQ914559.1  GI:6511075
KEYWORDS   GSS.
SOURCE     Oryza sativa (japonica cultivar-group)
            Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzoae; Oryza.
            1 (bases 1 to 814)
            Wing, R.A. and Dean, R.A.
            A BAC End Sequencing Framework to Sequence the Rice Genome
            Unpublished (1998)
            Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Seq primer: GGAACAGCTATGACCATG
            Class: BAC ends
            High quality sequence start: 21
            High quality sequence stop: 361.
FEATURES   source
            Location/Qualifiers
                1..814
                /organism="Oryza sativa (japonica cultivar-group)"
                /mol_type="genomic DNA"
                /cultivar="Nipponbare"
                /db_xref="taxon:39947"
                /clone="nhe0049W21r"
                /tissue_type="Leaf"
                /lab_host="E. coli DH10B"
                /clone_lib="CUGI Rice BAC Library (ECORI)"
                /notes="Vector: pBACIndigo; Site_1: EcorI; Site_2: EcorI;
                Rice is the most important food crop in the world. Half of
                the world population, especially those inhabiting highly
                populated areas of the humid tropics and subtropics, rely
                on rice as their primary source of carbohydrate.
                Monocotyledonous rice is a diploid plant (2n=24) with a
                haploid genome equivalent of 431 Mbp (Arumuganathan and

```

ORIGIN

RESULT 15

Genome Sequencing Center
Washington University School of Medicine

Washington University School of Medicine
Email: subhansingh@wustl.edu

Email: submissions@watson.wustl.edu

Plate: jnr57 row: d column: 03

Seq primer: -21UPpOT forward

Class: shotgun

High quality sequence start: 29

High quality sequence stop: 94.

FEATURES

source 1. .675

/organism="Brassica oleracea"

/mol type="genomic DNA

```
/db xref="taxon:3712"
```

```
/clone lib="B.oleracea001"
```

/notes="Vector: POTw13; Whole genome shotgun library from flowering buds. DNA was purified from a crude nuclear prep using Braseia oleracea Tol000DH3 buds provided by Thomas Osborn at the University of Wisconsin. Genomic DNA was provided by Pablo Rabinowicz (CSHL) and the shotgun library prepared at Washington University Genome Sequencing Center."

ORIGIN

Query Match 53.0%; Score 53; DB 8; Length 675;
Best Local Similarity 75.6%; Pred. No. 7.6e-08;

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:39:07 ; Search time 756.618 Seconds
(without alignments)
6468.225 Million cell updates/sec

Title: US-09-482-682-47_COPY_3944_4044

Perfect score: 101

Sequence: 1 aagcggcgtcgagtctag.....gtttgccctcccccggtgcc 101

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hg.*

3: gb_in.*

4: gb_on.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	6050	6	AX195205 Sequence
2	101	100.0	6567	6	AX128350 Sequence
3	101	100.0	6623	6	AX128345 Sequence
4	101	100.0	6639	6	AX128351 Sequence
5	101	100.0	6649	6	AX180726 Sequence
6	101	100.0	6695	6	AX128347 Sequence
7	101	100.0	6695	6	AX128353 Sequence
8	101	100.0	6695	6	AX128354 Sequence
9	101	100.0	6746	6	AX128344 Sequence
10	101	100.0	6818	6	AX128346 Sequence
11	101	100.0	6828	6	AX128340 Sequence
12	101	100.0	6833	6	AX128349 Sequence
13	101	100.0	6900	6	AX128341 Sequence
14	101	100.0	6956	6	AX128348 Sequence
15	101	100.0	7038	6	AX128342 Sequence
16	101	100.0	7960	6	BD268233 Adenoviru
17	101	100.0	7989	6	BD268236 Adenoviru
18	100	99.0	7231	6	BD268252 Adenoviru
19	99	98.0	280	6	CQ788637 Sequence

20	99	98.0	4291	12	AF302189	AF302189 Synthetic
21	99	98.0	5082	6	A91754	A91754 Sequence 10
22	99	98.0	5082	6	BD085110	BD085110 Vertebrat
23	99	98.0	5432	6	BD234590	BD234590 Screening
24	99	98.0	5432	6	AX026821	AX026821 Sequence
25	99	98.0	5650	6	AX226281	AX226281 Sequence
26	99	98.0	5731	6	AX202478	AX202478 Sequence
27	99	98.0	5771	12	AF060226	AF060226 Eukaryoti
28	99	98.0	6082	6	AR278592	AR278592 Sequence
29	99	98.0	6082	6	AR367288	AR367288 Sequence
30	99	98.0	6082	6	AR400324	AR400324 Sequence
31	99	98.0	6082	6	AR405591	AR405591 Sequence
32	99	98.0	6082	6	AR563971	AR563971 Sequence
33	99	98.0	6082	6	AX141045	AX141045 Sequence
34	99	98.0	6082	6	AX200905	AX200905 Sequence
35	99	98.0	6082	6	AX267561	AX267561 Sequence
36	99	98.0	6365	6	AX5113181	AX5113181 Sequence
37	99	98.0	6759	6	CQ788635	CQ788635 Sequence
38	99	98.0	6759	6	CQ788642	CQ788642 Sequence
39	99	98.0	6801	6	AX128352	AX128352 Sequence
40	99	98.0	6801	6	AX128355	AX128355 Sequence
41	99	98.0	7108	6	E36262	E36262 Human semap
42	99	98.0	7108	6	AX001326	AX001326 Sequence
43	99	98.0	7427	6	CQ768745	CQ768745 Sequence
44	99	98.0	7493	6	CQ768742	CQ768742 Sequence
45	99	98.0	8578	6	AR409005	AR409005 Sequence

ALIGNMENTS

RESULT 1
LOCUS AX195205 6050 bp DNA linear PAT 28-AUG-2001
DEFINITION Sequence 10 from Patent WO0151626.
ACCESSION AX195205
VERSION AX195205.1 GI:15385768
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Lu, X., Sun, L. and Zhang, Y.
TITLE Novel plasmid dna vectors
JOURNAL Patent: WO 0151626-A 10 19-JUL-2001;
ELIM BIOPHARMACEUTICALS, INC. (US)
FEATURES
source
1. 6050
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="commercial plasmid"

ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 6050;
Best Local Similarity 100.0%; Pred. No. 3.3e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AAGCGCGGTCGAGTCTAGAGGCGCCGTTTAAACCCGCTGATCAGCTCGCTGCT 60
Db 3828 AAGCGCGGTCGAGTCTAGAGGCGCCGTTTAAACCCGCTGATCAGCTCGCTGCT 3887
Qy 61 TCTAGTGTCCAGCCATCTGTTGTGCGCCCTCCCGCGTGC 101
Db 3888 TCTAGTGTCCAGCCATCTGTTGTGCGCCCTCCCGCGTGC 3928
RESULT 2
AX128350
LOCUS AX128350 6567 bp DNA linear PAT 15-MAY-2001
DEFINITION Sequence 11 from Patent WO0130843.
ACCESSION AX128350
VERSION AX128350.1 GI:14134863

KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.

REFERENCE
1
AUTHORS Barbas,C.F., Kadan,M. and Beerli,R.
TITLE Ligand activated transcriptional regulator proteins
JOURNAL Patent: WO 0130843-A 11 03-MAY-2001;
Novartis AG (CH); The Scripps Research Institute (US)

FEATURES
Location/Qualifiers
source
1..6567
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Construct E2CLBDAS"

ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 6567;
Best Local Similarity 100.0%; Pred. No. 3.3e-18; Mismatches 0; Indels 0; Gaps 0;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGCGCGCTCGAGTCTAGAGGCGCGGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 60
|||||
Db 2171 AAGCGCGCTCGAGTCTAGAGGCGCGGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 2230
|||||
Qy 61 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGCGTGCC 101
|||||
Db 2231 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGCGTGCC 2271
|||||

RESULT 3
AX128345
LOCUS AX128345 6623 bp DNA linear PAT 15-MAY-2001
DEFINITION Sequence 6 from Patent WO0130843.
ACCESSION AX128345
VERSION AX128345.1 GI:14134858
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.

REFERENCE
1
AUTHORS Barbas,C.F., Kadan,M. and Beerli,R.
TITLE Ligand activated transcriptional regulator proteins
JOURNAL Patent: WO 0130843-A 6 03-MAY-2001;
Novartis AG (CH); The Scripps Research Institute (US)

FEATURES
Location/Qualifiers
source
1..6623
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Construct C7LEBAS"

ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 6623;
Best Local Similarity 100.0%; Pred. No. 3.3e-18; Mismatches 0; Indels 0; Gaps 0;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGCGCGCTCGAGTCTAGAGGCGCGGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 60
|||||
Db 2227 AAGCGCGCTCGAGTCTAGAGGCGCGGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 2286
|||||
Qy 61 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGCGTGCC 101
|||||
Db 2287 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGCGTGCC 2327
|||||

RESULT 4
AX128351
LOCUS AX128351 6639 bp DNA linear PAT 15-MAY-2001
DEFINITION Sequence 12 from Patent WO0130843.
ACCESSION AX128351
VERSION AX128351.1 GI:14134864
KEYWORDS

SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.

REFERENCE
1
AUTHORS Barbas,C.F., Kadan,M. and Beerli,R.
TITLE Ligand activated transcriptional regulator proteins
JOURNAL Patent: WO 0130843-A 12 03-MAY-2001;
Novartis AG (CH); The Scripps Research Institute (US)

FEATURES
Location/Qualifiers
source
1..6639
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Construct E2CLBDBS"

ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 6639;
Best Local Similarity 100.0%; Pred. No. 3.3e-18; Mismatches 0; Indels 0; Gaps 0;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGCGCGCTCGAGTCTAGAGGCGCGGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 60
|||||
Db 2243 AAGCGCGCTCGAGTCTAGAGGCGCGGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 2302
|||||
Qy 61 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGCGTGCC 101
|||||
Db 2303 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGCGTGCC 2343
|||||

RESULT 5
AX180726
LOCUS AX180726 6649 bp DNA linear PAT 06-AUG-2001
DEFINITION Sequence 6 from Patent WO0146694.
ACCESSION AX180726
VERSION AX180726.1 GI:15132581
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.

REFERENCE
1
AUTHORS Joly,E.
TITLE A bioluminescence resonance energy transfer (bret) fusion molecule
and method of use
JOURNAL Patent: WO 0146694-A 6 28-JUN-2001;
Biosignal Packard Inc. (CA)

FEATURES
Location/Qualifiers
source
1..6649
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="DNA sequence for Rluc-PKA-EYFP construct"

ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 6649;
Best Local Similarity 100.0%; Pred. No. 3.3e-18; Mismatches 0; Indels 0; Gaps 0;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGCGCGCTCGAGTCTAGAGGCGCGGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 60
|||||
Db 2611 AAGCGCGCTCGAGTCTAGAGGCGCGGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 2670
|||||
Qy 61 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGCGTGCC 101
|||||
Db 2671 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGCGTGCC 2711
|||||

RESULT 6
AX128347
LOCUS AX128347 6695 bp DNA linear PAT 15-MAY-2001
DEFINITION Sequence 8 from Patent WO0130843.
ACCESSION AX128347
VERSION AX128347.1 GI:14134860
KEYWORDS

SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.

REFERENCE 1
AUTHORS Barbas, C.F., Kadan, M. and Beerli, R.
TITLE Ligand activated transcriptional regulator proteins
JOURNAL Patent: WO 0130843-A 8 03-MAY-2001;
Novartis AG (CH); The Scripps Research Institute (US)

FEATURES
source
1..6695
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Construct C7LBD8S"

ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 6695;
Best Local Similarity 100.0%; Pred. No. 3.3e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGCGCGCTCGAGTCTAGAGGCGCGCTTTAAACCCGCTGATCAGCCTCGACTGTGCCT 60
|||||
Db 2299 AAGCGCGCTCGAGTCTAGAGGCGCGCTTTAAACCCGCTGATCAGCCTCGACTGTGCCT 2358
|||||

Qy 61 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGCGGCC 101
|||||
Db 2359 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGCGGCC 2399
|||||

RESULT 7
AX128353
LOCUS AX128353 6695 bp DNA linear PAT 15-MAY-2001
DEFINITION Sequence 14 from Patent WO0130843.
ACCESSION AX128353
VERSION AX128353.1 GI:14134866
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.

REFERENCE 1
AUTHORS Barbas, C.F., Kadan, M. and Beerli, R.
TITLE Ligand activated transcriptional regulator proteins
JOURNAL Patent: WO 0130843-A 14 03-MAY-2001;
Novartis AG (CH); The Scripps Research Institute (US)

FEATURES
source
1..6695
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Construct LBDBSG400V"

ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 6695;
Best Local Similarity 100.0%; Pred. No. 3.3e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGCGCGCTCGAGTCTAGAGGCGCGCTTTAAACCCGCTGATCAGCCTCGACTGTGCCT 60
|||||
Db 2299 AAGCGCGCTCGAGTCTAGAGGCGCGCTTTAAACCCGCTGATCAGCCTCGACTGTGCCT 2358
|||||

Qy 61 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGCGGCC 101
|||||
Db 2359 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGCGGCC 2399
|||||

RESULT 8
AX128354
LOCUS AX128354 6695 bp DNA linear PAT 15-MAY-2001
DEFINITION Sequence 15 from Patent WO0130843.
ACCESSION AX128354
VERSION AX128354.1 GI:14134867
KEYWORDS
SOURCE synthetic construct

ORGANISM synthetic construct
other sequences; artificial sequences.

REFERENCE 1
AUTHORS Barbas, C.F., Kadan, M. and Beerli, R.
TITLE Ligand activated transcriptional regulator proteins
JOURNAL Patent: WO 0130843-A 15 03-MAY-2001;
Novartis AG (CH); The Scripps Research Institute (US)

FEATURES
source
1..6695
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Construct LBDBSG521R"

ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 6695;
Best Local Similarity 100.0%; Pred. No. 3.3e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGCGCGCTCGAGTCTAGAGGCGCGCTTTAAACCCGCTGATCAGCCTCGACTGTGCCT 60
|||||
Db 2299 AAGCGCGCTCGAGTCTAGAGGCGCGCTTTAAACCCGCTGATCAGCCTCGACTGTGCCT 2358
|||||

Qy 61 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGCGGCC 101
|||||
Db 2359 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGCGGCC 2399
|||||

RESULT 9
AX128344
LOCUS AX128344 6746 bp DNA linear PAT 15-MAY-2001
DEFINITION Sequence 5 from Patent WO0130843.
ACCESSION AX128344
VERSION AX128344.1 GI:14134857
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.

REFERENCE 1
AUTHORS Barbas, C.F., Kadan, M. and Beerli, R.
TITLE Ligand activated transcriptional regulator proteins
JOURNAL Patent: WO 0130843-A 5 03-MAY-2001;
Novartis AG (CH); The Scripps Research Institute (US)

FEATURES
source
1..6746
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Construct C7LBDAL"

ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 6746;
Best Local Similarity 100.0%; Pred. No. 3.3e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGCGCGCTCGAGTCTAGAGGCGCGCTTTAAACCCGCTGATCAGCCTCGACTGTGCCT 60
|||||
Db 2350 AAGCGCGCTCGAGTCTAGAGGCGCGCTTTAAACCCGCTGATCAGCCTCGACTGTGCCT 2409
|||||

Qy 61 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGCGGCC 101
|||||
Db 2410 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGCGGCC 2450
|||||

RESULT 10
AX128346
LOCUS AX128346 6818 bp DNA linear PAT 15-MAY-2001
DEFINITION Sequence 7 from Patent WO0130843.
ACCESSION AX128346
VERSION AX128346.1 GI:14134859
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct

```
other sequences; artificial sequences.
REFERENCE 1
AUTHORS Barbas,C.F., Kadan,M. and Beerli,R.
TITLE Ligand activated transcriptional regulator proteins
JOURNAL Patent: WO 0130843-A 7 03-MAY-2001;
Novartis AG (CH); The Scripps Research Institute (US)
FEATURES
source
1..6818
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Construct C7LBDAS"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 6818;
Best Local Similarity 100.0%; Pred. No. 3.2e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAGCGCGCTCGAGTCTAGAGGCGCGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 60
Db 2422 AAGCGCGCTCGAGTCTAGAGGCGCGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 2481
QY 61 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGGTGCC 101
Db 2482 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGGTGCC 2522
RESULT 11
AX128340
LOCUS AX128340 6828 bp DNA linear PAT 15-MAY-2001
DEFINITION Sequence 1 from Patent WO0130843.
ACCESSION AX128340
VERSION AX128340.1 GI:14134853
KEYWORDS
SOURCE
synthetic construct
ORGANISM
synthetic construct
other sequences; artificial sequences.
REFERENCE 1
AUTHORS Barbas,C.F., Kadan,M. and Beerli,R.
TITLE Ligand activated transcriptional regulator proteins
JOURNAL Patent: WO 0130843-A 1 03-MAY-2001;
Novartis AG (CH); The Scripps Research Institute (US)
FEATURES
source
1..6828
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Construct C7LBDAS"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 6828;
Best Local Similarity 100.0%; Pred. No. 3.2e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAGCGCGCTCGAGTCTAGAGGCGCGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 60
Db 2432 AAGCGCGCTCGAGTCTAGAGGCGCGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 2491
QY 61 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGGTGCC 101
Db 2492 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGGTGCC 2532
RESULT 12
AX128349
LOCUS AX128349 6833 bp DNA linear PAT 15-MAY-2001
DEFINITION Sequence 10 from Patent WO0130843.
ACCESSION AX128349
VERSION AX128349.1 GI:14134862
KEYWORDS
SOURCE
synthetic construct
ORGANISM
synthetic construct
other sequences; artificial sequences.
other sequences; artificial sequences.
REFERENCE 1
AUTHORS Barbas,C.F., Kadan,M. and Beerli,R.
TITLE Ligand activated transcriptional regulator proteins
JOURNAL Patent: WO 0130843-A 7 03-MAY-2001;
Novartis AG (CH); The Scripps Research Institute (US)
FEATURES
source
1..6818
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Construct C7LBDAS"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 6818;
Best Local Similarity 100.0%; Pred. No. 3.2e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAGCGCGCTCGAGTCTAGAGGCGCGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 60
Db 2422 AAGCGCGCTCGAGTCTAGAGGCGCGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 2481
QY 61 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGGTGCC 101
Db 2482 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGGTGCC 2522
RESULT 11
AX128340
LOCUS AX128340 6828 bp DNA linear PAT 15-MAY-2001
DEFINITION Sequence 1 from Patent WO0130843.
ACCESSION AX128340
VERSION AX128340.1 GI:14134853
KEYWORDS
SOURCE
synthetic construct
ORGANISM
synthetic construct
other sequences; artificial sequences.
REFERENCE 1
AUTHORS Barbas,C.F., Kadan,M. and Beerli,R.
TITLE Ligand activated transcriptional regulator proteins
JOURNAL Patent: WO 0130843-A 1 03-MAY-2001;
Novartis AG (CH); The Scripps Research Institute (US)
FEATURES
source
1..6828
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Construct C7LBDAS"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 6828;
Best Local Similarity 100.0%; Pred. No. 3.2e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAGCGCGCTCGAGTCTAGAGGCGCGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 60
Db 2432 AAGCGCGCTCGAGTCTAGAGGCGCGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 2491
QY 61 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGGTGCC 101
Db 2492 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGGTGCC 2532
RESULT 12
AX128349
LOCUS AX128349 6833 bp DNA linear PAT 15-MAY-2001
DEFINITION Sequence 10 from Patent WO0130843.
ACCESSION AX128349
VERSION AX128349.1 GI:14134862
KEYWORDS
SOURCE
synthetic construct
ORGANISM
synthetic construct
other sequences; artificial sequences.
other sequences; artificial sequences.
REFERENCE 1
AUTHORS Barbas,C.F., Kadan,M. and Beerli,R.
TITLE Ligand activated transcriptional regulator proteins
JOURNAL Patent: WO 0130843-A 10 03-MAY-2001;
Novartis AG (CH); The Scripps Research Institute (US)
FEATURES
source
1..6833
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Construct C7LBDCS"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 6833;
Best Local Similarity 100.0%; Pred. No. 3.2e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAGCGCGCTCGAGTCTAGAGGCGCGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 60
Db 2437 AAGCGCGCTCGAGTCTAGAGGCGCGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 2496
QY 61 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGGTGCC 101
Db 2497 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGGTGCC 2537
RESULT 13
AX128341
LOCUS AX128341 6900 bp DNA linear PAT 15-MAY-2001
DEFINITION Sequence 2 from Patent WO0130843.
ACCESSION AX128341
VERSION AX128341.1 GI:14134854
KEYWORDS
SOURCE
synthetic construct
ORGANISM
synthetic construct
other sequences; artificial sequences.
REFERENCE 1
AUTHORS Barbas,C.F., Kadan,M. and Beerli,R.
TITLE Ligand activated transcriptional regulator proteins
JOURNAL Patent: WO 0130843-A 2 03-MAY-2001;
Novartis AG (CH); The Scripps Research Institute (US)
FEATURES
source
1..6900
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Construct 2C7LBDBS"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 6900;
Best Local Similarity 100.0%; Pred. No. 3.2e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAGCGCGCTCGAGTCTAGAGGCGCGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 60
Db 2504 AAGCGCGCTCGAGTCTAGAGGCGCGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 2563
QY 61 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGGTGCC 101
Db 2564 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGGTGCC 2604
RESULT 14
AX128348
LOCUS AX128348 6956 bp DNA linear PAT 15-MAY-2001
DEFINITION Sequence 9 from Patent WO0130843.
ACCESSION AX128348
VERSION AX128348.1 GI:14134861
KEYWORDS
SOURCE
synthetic construct
ORGANISM
synthetic construct
other sequences; artificial sequences.
other sequences; artificial sequences.
REFERENCE 1
```

AUTHORS Barbas,C.F., Kadan,M. and Beerli,R.
TITLE Ligand activated transcriptional regulator proteins
JOURNAL Patent: WO 0130843-A 9 03-MAY-2001;
Novartis AG (CH) ; The Scripps Research Institute (US)
FEATURES
source 1..6956
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Construct C7LBDCL"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 6956;
Best Local Similarity 100.0%; Pred.No.3.2e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AAGCGCGCGCTCGAGTCTAGAGGGCGCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCT 60
|||||
Db 2560 AAGCGCGCGCTCGAGTCTAGAGGGCGCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCT 2619
|||||
Qy 61 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGCGTGCC 101
|||||
Db 2620 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGCGTGCC 2660
|||||
RESULT 15
AX128342
LOCUS AX128342 7038 bp DNA linear PAT 15-MAY-2001
DEFINITION Sequence 3 from Patent WO0130843.
ACCESSION AX128342
VERSION AX128342.1 GI:14134855
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE 1
AUTHORS Barbas,C.F., Kadan,M. and Beerli,R.
TITLE Ligand activated transcriptional regulator proteins
JOURNAL Patent: WO 0130843-A 3 03-MAY-2001;
Novartis AG (CH) ; The Scripps Research Institute (US)
FEATURES
source 1..7038
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Construct 2C7LBDCL"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 7038;
Best Local Similarity 100.0%; Pred.No.3.2e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AAGCGCGCGCTCGAGTCTAGAGGGCGCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCT 60
|||||
Db 2642 AAGCGCGCGCTCGAGTCTAGAGGGCGCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCT 2701
|||||
Qy 61 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGCGTGCC 101
|||||
Db 2702 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGCGTGCC 2742
|||||

Search completed: July 14, 2005, 14:03:32
Job time : 756.618 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:35:42 ; Search time 142.398 Seconds
(without alignments)
4198.742 Million cell updates/sec

Title: US-09-482-682-47_COPY_3944_4044

Perfect score: 101

Sequence: 1 aagcgccgtcagctag.....gtttgccccccccctgtgcc 101

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002as:*

7: Geneseqn2002bs:*

8: Geneseqn2003as:*

9: Geneseqn2003bs:*

10: Geneseqn2003cs:*

11: Geneseqn2003ds:*

12: Geneseqn2004as:*

13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	2017	8 ADA05887	Ada05887 Human NOV
2	101	100.0	2017	12 ADNG3050	Adnc3050 Human NOV
3	101	100.0	2017	12 ADO42504	Ado42504 Human NOV
4	101	100.0	2017	12 ADO42500	Ado42500 Human NOV
5	101	100.0	2022	12 ADO42506	Ado42506 Human NOV
6	101	100.0	6050	5 AAD10237	Aad10237 Commercia
7	101	100.0	6567	4 AAD06054	Aad06054 Plasmid E
8	101	100.0	6623	4 AAD06049	Aad06049 Plasmid C
9	101	100.0	6639	4 AAD06055	Aad06055 Plasmid E
10	101	100.0	6695	4 AAD06051	Aad06051 Plasmid C
11	101	100.0	6695	4 AAD06057	Aad06057 Plasmid C
12	101	100.0	6695	4 AAD06058	Aad06058 Plasmid C
13	101	100.0	6746	4 AAD06048	Aad06048 Plasmid C
14	101	100.0	6818	4 AAD06050	Aad06050 Plasmid C
15	101	100.0	6828	4 AAD06044	Aad06044 Plasmid 2
16	101	100.0	6833	4 AAD06053	Aad06053 Plasmid C
17	101	100.0	6900	4 AAD06045	Aad06045 Plasmid 2
18	101	100.0	6956	4 AAD06052	Aad06052 Plasmid C
19	101	100.0	7038	4 AAD06046	Aad06046 Plasmid 2
20	101	100.0	7586	6 ABA92644	Ab92644 Cholera t

21	101	100.0	7960	3 AAA59072	Aaa59072 Nucleotid
22	101	100.0	7960	6 ABA94274	Ab94274 Nucleotid
23	101	100.0	7960	10 ADB75120	Adb75120 Plasmid p
24	101	100.0	7960	10 ADF48754	Adf48754 Fibre exp
25	101	100.0	7989	3 AAA59075	Aaa59075 Nucleotid
26	101	100.0	7989	6 ABA94277	Ab94277 Nucleotid
27	101	100.0	7989	10 ADB75123	Adb75123 Plasmid p
28	101	100.0	7989	10 ADF48757	Adf48757 Fibre exp
29	101	100.0	9626	12 ADM97793	Adm97793 Plasmid p
30	100	99.0	1353	10 ADC26320	Adc26320 Human NOV
31	100	99.0	1353	12 ADM35637	Adm35637 Novel hum
32	100	99.0	1353	12 ADO42484	Ado42484 Human NOV
33	100	99.0	1420	8 ACC62251	Acc62251 Human NOV
34	100	99.0	1461	8 ACC62237	Acc62237 Human NOV
35	100	99.0	1770	10 ADJ94793	Adj94793 Novel NOV
36	100	99.0	1772	10 ADJ94791	Adj94791 Novel NOV
37	100	99.0	1772	10 ADJ94795	Adj94795 Novel NOV
38	100	99.0	1822	10 ACD19336	Acd19336 cDNA enco
39	100	99.0	1822	10 ACD19334	Acd19334 cDNA enco
40	100	99.0	6498	10 AAD62465	Aad62465 Human MCH
41	100	99.0	6498	10 ADH53343	Adh53343 pCDNA3 P1
42	100	99.0	7231	3 AAA59090	Aaa59090 Nucleotid
43	100	99.0	7231	6 ABA94286	Ab94286 Nucleotid
44	100	99.0	7231	10 ADB75132	Adb75132 Plasmid p
45	100	99.0	7231	10 ADF48774	Adf48774 Adenoviru

ALIGNMENTS

RESULT 1

ADA05887

ID ADA05887 standard; cDNA; 2017 BP.

XX AC ADA05887;

XX AC ADA05887;

XX AC ADA05887;

DT 06-NOV-2003 (first entry)

XX DE

Human NOV55b encoding cDNA SEQ ID NO:247.

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

PA (LILL//) LI L.
PA (GUOX//) GUO X.
PA (PATT//) PATTURAJAN M.
PA (SPYT//) SPYTEK K A.
PA (EDIN//) EDINGER S R.
PA (ELLE//) ELLERMAN K.
PA (MALY//) MALYANKAR U M.
PA (ORTT//) ORT T.
PA (GORM//) GORMAN L.
PA (ZERH//) ZERHUSEN B D.
PA (ANDE//) ANDERSON D W.
PA (ZHON//) ZHONG M.
PA (CATT//) CATTERTON E.
PA (JIWV//) JI W.
PA (MILL//) MILLER C E.
PA (RAST//) RASTELLI L.
PA (STON//) STONE D J.
PA (PENA//) PENA C E A.
PA (SHEN//) SHENOY S G.
PA (SHIM//) SHIMKETS R A.
PA (ROTH//) ROTHENBERG M E.
PA (LEAC//) LEACH M D.
PA (AGEE//) AGEE M L.
PA (BERG//) BERGHS C.
PA (DIPV//) DIPIPPO V A.
PA (EISE//) EISEN A.
PA (GANG//) GANGOLLI E A.
PA (RIEG//) RIEGER D K.
PA (SPAD//) SPADERNA S K.
XX
XX Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;
PI Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;
PI Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;
PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;
PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;
PI Eisen A, Gangolli EA, Rieger DK, Spaderna SK;
XX
XX WPI; 2004-213931/20.
DR P-PSDB; ADN63051.
XX
XX Isolated NOVX polypeptides and nucleic acids, useful for preventing,
PT diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.
XX
XX Claim 20; SEQ ID NO 247; 395pp; English.
XX
XX The invention relates to isolated NOVX polypeptides and polynucleotides.
CC NOVX polypeptides and polynucleotides are used to prevent, diagnose or
CC treat a medical condition in human related to the aberrant expression and
CC activity of NOVX polypeptides. For example, NOVX polypeptides and
CC polynucleotides may be used to treat disorders associated with decreased
CC expression or activity of NOVX by supplementing the patient our
CC production or to rectify mutations. Conversely, antisense NA molecules
CC may be administered to down regulate expression of NOVX polypeptides by
CC binding with the cells own genes and preventing their expression. NOVX
CC polynucleotides and complementary sequences may also be used as DNA
CC probes in diagnostic assays to detect and quantitate the presence of
CC similar sequences in samples, and so which patients may be in need of
CC restorative therapy. NOVX polypeptides may also be used as antigens in
CC the production of antibodies and in assays to identify modulators
CC (agonists and antagonists) of the expression and activity of NOVX. The
CC anti-NOVX polypeptide antibodies, agonists and antagonists may also be
CC used to modulate NOVX polynucleotide expression and activity of NOVX
CC polypeptides. The anti-NOVX polypeptide antibodies may also be used as
CC diagnostic agents for detecting the presence of NOVX in samples. NOVX
CC polypeptides and polynucleotides may be used in this way to prevent,
CC diagnose and treat: metabolic disorders, diabetes, obesity, infectious
CC disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative
CC disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders,
CC haematopoietic disorders, and the various dyslipidaemias, metabolic
CC disturbances associated with obesity, the metabolic syndrome X and
CC wasting disorders associated with chronic diseases and various cancers.
CC They may also be used as antibacterial agents. The present sequence
CC represents DNA encoding a human NOVX protein.

XX SQ Sequence 2017 BP; 351 A; 625 C; 599 G; 441 T; 0 U; 1 Other;
Query Match 100.0%; Score 101; DB 12; Length 2017;
Best Local Similarity 100.0%; Pred. No. 2.4e-24;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AAGCGCGCTCGAGTCTAGAGGCGCGCTTTAAACCCGCTGATCAGCTCGACTGTGCCT 60
Db 1839 AAGCGCGCGCTCGAGTCTAGAGGCGCGCTTTAAACCCGCTGATCAGCTCGACTGTGCCT 1898
Qy 61 TCTAGTTGCCAGCCATCTGTGTTGCCCTCCCGCGTGC 101
Db 1899 TCTAGTTGCCAGCCATCTGTGTTGCCCTCCCGCGTGC 1939
RESULT 3
ID ADO42504 standard; cDNA; 2017 BP.
XX
XX ADO42504;
XX AC
XX DT 15-JUL-2004 (first entry)
XX DE Human NOVX polynucleotide #177.
XX KW Human; NOVX; gene; ss; cancer; atherosclerosis; diabetes;
KW Alzheimer's disease; Parkinson's disease; graft-versus-host disease;
KW scleroderma; hypertension; haemophilia;
KW idiopathic thrombocytopenic purpura; immunodeficiency; AIDS;
KW dyslipidemia; obesity; Crohn's disease; bronchial asthma; anorexia;
KW cancer-associated cachexia; multiple sclerosis; fertility.
XX OS Homo sapiens.
XX FN US2004058338-A1.
XX PD 25-MAR-2004.
XX PF 02-DEC-2002; 2002US-00307817.
XX PR 03-DEC-2001; 2001US-0336881P.
XX PR 05-DEC-2001; 2001US-0336820P.
XX PR 07-DEC-2001; 2001US-0338285P.
XX PR 07-DEC-2001; 2001US-0338318P.
XX PR 10-DEC-2001; 2001US-0338989P.
XX PR 10-DEC-2001; 2001US-0339022P.
XX PR 11-DEC-2001; 2001US-0339314P.
XX PR 11-DEC-2001; 2001US-0339516P.
XX PR 11-DEC-2001; 2001US-0339517P.
XX PR 11-DEC-2001; 2001US-0339611P.
XX PR 12-DEC-2001; 2001US-0340981P.
XX PR 12-DEC-2001; 2001US-0341346P.
XX PR 14-DEC-2001; 2001US-0340390P.
XX PR 14-DEC-2001; 2001US-0340440P.
XX PR 14-DEC-2001; 2001US-0340565P.
XX PR 14-DEC-2001; 2001US-0340608P.
XX PR 14-DEC-2001; 2001US-0341144P.
XX PR 17-DEC-2001; 2001US-0341477P.
XX PR 17-DEC-2001; 2001US-0341540P.
XX PR 18-DEC-2001; 2001US-0341768P.
XX PR 20-DEC-2001; 2001US-0342592P.
XX PR 31-DEC-2001; 2001US-0344903P.
XX PR 01-FEB-2002; 2002US-0353286P.
XX PR 01-FEB-2002; 2002US-0353288P.
XX PR 26-FEB-2002; 2002US-0359599P.
XX PR 26-FEB-2002; 2002US-0359626P.
XX PR 26-FEB-2002; 2002US-0359671P.
XX PR 27-FEB-2002; 2002US-0359914P.
XX PR 27-FEB-2002; 2002US-0359956P.
XX PR 28-FEB-2002; 2002US-0360924P.
XX PR 28-FEB-2002; 2002US-0360964P.
XX PR 28-FEB-2002; 2002US-0361028P.

PR 28-FEB-2002; 2002US-0361256P.
 PR 28-FEB-2002; 2002US-0361264P.
 PR 05-MAR-2002; 2002US-0361770P.
 PR 05-MAR-2002; 2002US-0362230P.
 PR 13-MAR-2002; 2002US-0364181P.
 PR 13-MAR-2002; 2002US-0364238P.
 PR 15-MAR-2002; 2002US-0364978P.
 PR 15-MAR-2002; 2002US-0365025P.
 PR 17-APR-2002; 2002US-0373288P.
 PR 15-MAY-2002; 2002US-0380981P.
 PR 16-MAY-2002; 2002US-0381004P.
 PR 17-MAY-2002; 2002US-0381495P.
 PR 28-MAY-2002; 2002US-0383534P.
 PR 28-MAY-2002; 2002US-0383744P.
 PR 29-MAY-2002; 2002US-0383829P.
 PR 29-MAY-2002; 2002US-0384024P.
 PR 02-JUL-2002; 2002US-0393332P.
 PR 06-AUG-2002; 2002US-0401315P.
 PR 07-AUG-2002; 2002US-0401788P.
 PR 20-AUG-2002; 2002US-0404676P.
 PR 23-AUG-2002; 2002US-0405400P.
 PR 23-AUG-2002; 2002US-0405684P.
 PR 23-AUG-2002; 2002US-0405687P.
 PR 23-AUG-2002; 2002US-0405698P.
 PR 26-AUG-2002; 2002US-0406353P.
 XX (AGEE//) AGEE M L.
 PA (ALSO//) ALSOBROOK J P.
 PA (ANDE//) ANDERSON D W.
 PA (BERG//) BERGHS C.
 PA (BOLD//) BOLDOG F L.
 PA (BURG//) BURGESS C E.
 PA (CAIT//) CAITERTON E.
 PA (DIP//) DIPIPPO V A.
 PA (EDIN//) EDINGER S R.
 PA (EISE//) EISEN A.
 PA (ELLE//) ELLERMAN K.
 PA (GANG//) GANGOLLI E A.
 PA (GERL//) GERLACH V.
 PA (GORM//) GORMAN L.
 PA (ROTH//) ROTHBERG B G.
 PA (GUOX//) GUO X S.
 PA (HERR//) HERRMANN J L.
 PA (HALV//) HALVORSEN Y.
 PA (JIW//) JI W.
 PA (KEKU//) KEKUDA R.
 PA (KHRA//) KHRAMTSOV N V.
 PA (LARO//) LAROCHELLE W J.
 PA (LRPL//) LEPLEY D M.
 PA (LILL//) LI L.
 PA (MACD//) MACDOUGALL J R.
 PA (MILL//) MILLER C E.
 PA (ORTT//) ORT T.
 PA (PADI//) PADIGARU M.
 PA (PENA//) PENA C E A.
 PA (PEYM//) PEYMAN J A.
 PA (RIEG//) RIEGER D K.
 PA (ROTH//) ROTHENBERG M E.
 PA (SHEN//) SHENOY S G.
 PA (SMIT//) SMITHSON G.
 PA (SPAD//) SPADERNA S K.
 PA (SPYT//) SPYTEK K A.
 PA (STON//) STONE D J.
 PA (TAUP//) TAUPIER R J.
 PA (VERN//) VERNET C A M.
 PA (VOSS//) VOSS E Z.
 PA (ZHON//) ZHONG M.
 XX
 PI Agee ML, Alsbrook JP, Anderson DW, Berghs C, Boldog FL;
 PI Burgess CE, Catterton E, Dipippo VA, Edinger SR, Eisen A;
 PI Ellerman K, Gangolli EA, Gerlach V, Gorman L, Rothberg BG, Herrmann JL, Halvorsen Y, Ji W, Kekuda R, Khrantsov NV;
 PI

PI Larochele WJ, Lepley DM, Li L, Macdougall JR, Miller CE, Ort T;
 PI Padigar M, Patturajan M, Pena CEA, Peyman JA, Rieger DK;
 PI Stone DU, Taupier RJ, Vernet CAM, Voss EZ, Zhong M;
 XX WPI; 2004-268786/25.
 DR P-PSDB; ADO42505.
 XX
 PT New human NOVX polypeptides and nucleic acid molecules, useful for
 PT diagnosing, preventing or treating NOVX-associated disorder, e.g. cancer,
 PT atherosclerosis, diabetes, Alzheimer's disease, Parkinson's disease or
 PT scleroderma.
 XX
 PS Claim 20; SEQ ID NO 353; 610pp; English.
 XX
 CC The invention relates to human NOVX polypeptides and the polynucleotides
 CC encoding them. The invention also relates to antibodies specific to the
 CC NOVX polypeptides. The polypeptides, polynucleotides and antibodies are
 CC useful for manufacturing a medicament for treating a syndrome associated
 CC with a human disease, such as a pathology associated with the NOVX
 CC polypeptide. The sequences are useful for diagnosing, treating or
 CC preventing a NOVX-associated disorder, e.g., cancer, atherosclerosis,
 CC diabetes, Alzheimer's disease, Parkinson's disease, graft-versus-host
 CC disease, scleroderma, hypertension, haemophilia, idiopathic
 CC thrombocytopenic purpura, immunodeficiencies, AIDS, dyslipidemia,
 CC obesity, Crohn's disease, bronchial asthma, anorexia, cancer-associated
 CC cachexia, multiple sclerosis or fertility. The nucleic acids may be used
 CC as hybridisation probes, in chromosome mapping, in tissue typing, in
 CC preventive medicine or in pharmacogenomics. This sequence represents a
 CC human NOVX polynucleotide of the invention.
 XX
 SQ Sequence 2017 BP; 345 A; 613 C; 617 G; 442 T; 0 U; 0 Other;
 Query Match 100.0%; Score 101; DB 12; Length 2017;
 Best Local Similarity 100.0%; Pred. No. 2.4e-24;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AAGCGCGCGCTCGAGTCTAGAGGGCGCGGTTAAACCGCTGATCAGCTCGACTGTCCT 60
 DB 1839 AAGCGCGCGCTCGAGTCTAGAGGGCGCGGTTAAACCGCTGATCAGCTCGACTGTCCT 1898
 QY 61 TCTAGTTGCCAGCATCTGTTTGGCCCTCCCGCGTGCC 101
 DB 1899 TCTAGTTGCCAGCATCTGTTTGGCCCTCCCGCGTGCC 1939
 RESULT 4
 ID ADO42500 standard; cDNA; 2017 BP.
 XX ADO42500;
 AC ADO42500;
 XX
 DT 15-JUL-2004 (first entry)
 XX
 DE Human NOVX polynucleotide #175.
 XX
 KW Human; NOVX; gene; ss; cancer; atherosclerosis; diabetes;
 KW Alzheimer's disease; Parkinson's disease; graft-versus-host disease;
 KW scleroderma; hypertension; haemophilia;
 KW idiopathic thrombocytopenic purpura; immunodeficiency; AIDS;
 KW dyslipidemia; obesity; Crohn's disease; bronchial asthma; anorexia;
 KW cancer-associated cachexia; multiple sclerosis; fertility.
 XX
 OS Homo sapiens.
 OS
 XX US2004058338-A1.
 PN
 XX 25-MAR-2004.
 PD
 XX 02-DEC-2002; 2002US-00307817.
 PF
 XX 03-DEC-2001; 2001US-0336881P.
 PR
 XX 05-DEC-2001; 2001US-0336820P.
 PR


```
PR 07-DEC-2001; 2001US-0338285P.
PR 07-DEC-2001; 2001US-0338318P.
PR 10-DEC-2001; 2001US-0338989P.
PR 10-DEC-2001; 2001US-0339022P.
PR 11-DEC-2001; 2001US-0339314P.
PR 11-DEC-2001; 2001US-0339516P.
PR 11-DEC-2001; 2001US-0339517P.
PR 11-DEC-2001; 2001US-0339611P.
PR 12-DEC-2001; 2001US-0340981P.
PR 12-DEC-2001; 2001US-0341346P.
PR 14-DEC-2001; 2001US-0340390P.
PR 14-DEC-2001; 2001US-034040P.
PR 14-DEC-2001; 2001US-0340565P.
PR 14-DEC-2001; 2001US-0340608P.
PR 14-DEC-2001; 2001US-0341144P.
PR 17-DEC-2001; 2001US-0341477P.
PR 17-DEC-2001; 2001US-0341540P.
PR 18-DEC-2001; 2001US-0341768P.
PR 20-DEC-2001; 2001US-0344592P.
PR 31-DEC-2001; 2001US-0344903P.
PR 01-FEB-2002; 2002US-0353286P.
PR 01-FEB-2002; 2002US-0353288P.
PR 26-FEB-2002; 2002US-0359599P.
PR 26-FEB-2002; 2002US-0359626P.
PR 26-FEB-2002; 2002US-0359671P.
PR 27-FEB-2002; 2002US-0359914P.
PR 27-FEB-2002; 2002US-0359956P.
PR 28-FEB-2002; 2002US-0360924P.
PR 28-FEB-2002; 2002US-0360964P.
PR 28-FEB-2002; 2002US-0361028P.
PR 28-FEB-2002; 2002US-0361256P.
PR 28-FEB-2002; 2002US-0361264P.
PR 05-MAR-2002; 2002US-0361770P.
PR 05-MAR-2002; 2002US-0362230P.
PR 13-MAR-2002; 2002US-0364181P.
PR 13-MAR-2002; 2002US-0364238P.
PR 15-MAR-2002; 2002US-0364978P.
PR 15-MAR-2002; 2002US-0365025P.
PR 17-APR-2002; 2002US-0373288P.
PR 15-MAY-2002; 2002US-0380981P.
PR 16-MAY-2002; 2002US-0381004P.
PR 17-MAY-2002; 2002US-0381495P.
PR 28-MAY-2002; 2002US-0383534P.
PR 28-MAY-2002; 2002US-0383744P.
PR 29-MAY-2002; 2002US-0383829P.
PR 29-MAY-2002; 2002US-0384024P.
PR 02-JUL-2002; 2002US-0393332P.
PR 06-AUG-2002; 2002US-0401315P.
PR 07-AUG-2002; 2002US-0401789P.
PR 20-AUG-2002; 2002US-0404676P.
PR 23-AUG-2002; 2002US-0405400P.
PR 23-AUG-2002; 2002US-0405684P.
PR 23-AUG-2002; 2002US-0405687P.
PR 23-AUG-2002; 2002US-0405698P.
PR 26-AUG-2002; 2002US-0406353P.
XX (AGEE/) AGEE M L.
PA (ALSO/) ALSOBROOK J P.
PA (ANDE/) ANDERSON D W.
PA (BERG/) BERGHS C.
PA (BOLD/) BOLDOG F L.
PA (BURG/) BURGESS C E.
PA (CATI/) CATTERTON E.
PA (DIFI/) DIPPIO V A.
PA (EDIN/) EDINGER S R.
PA (EISE/) EISEN A.
PA (ELLE/) ELLERMAN K.
PA (GANG/) GANCOLLI E A.
PA (GERL/) GERLACH V.
PA (GORM/) GORMAN L.
PA (ROTH/) ROTHBERG B G.
PA (GUOX/) GUO X S.
PA (HERR/) HERRMANN J L.
PA (HALV/) HALVORSEN Y.
PA (JIWW/) JI W.
PA (KEKU/) KEKUDA R.
PA (KIRA/) KHRAMTSOV N V.
PA (LABO/) LAROUCHELLE W J.
PA (LEPL/) LEFLEY D M.
PA (LILL/) LI L.
PA (MACD/) MACDOUGALL J R.
PA (MILL/) MILLER C E.
PA (ORTT/) ORT T.
PA (PADI/) PADIGARU M.
PA (PATT/) PATTURAJAN M.
PA (PENA/) PENNA C E A.
PA (PEYM/) PEYMAN J A.
PA (RIEG/) RIEGER D K.
PA (ROTH/) ROTHENBERG M E.
PA (SHEN/) SHENOY S G.
PA (SMIT/) SMITHSON G.
PA (SPAD/) SPADERNA S K.
PA (SPYT/) SPYTEK K A.
PA (STON/) STONE D J.
PA (TAUP/) TAUPIER R J.
PA (VERN/) VERNET C A M.
PA (VOSS/) VOSS E Z.
PA (ZHON/) ZHONG M.
XX
PI Agee ML, Alsobrook JP, Anderson DW, Berghs C, Boldog FL,
PI Burgess CE, Catterton E, Dipippo VA, Edinger SR, Eisen A,
PI Ellerman K, Gangolli EA, Gerlach V, Gorman L, Rothberg BG, Guo XS,
PI Herrmann JL, Halvorsen Y, Ji W, Kekuda R, Khrantsov NV,
PI Larochele WJ, Lepley DM, Li L, Macdougall JR, Miller CE, Ort T,
PI Padigaru M, Patturajan M, Pena CE, Peyman JA, Rieger DK,
PI Rothenberg ME, Shenoy SG, Smithson G, Spaderna SK, Spytek KA,
PI Stone DJ, Taupier RJ, Vernet CAM, Voss EZ, Zhong M;
XX
XX WPI; 2004-268786/25.
DR P-PSDB; ADO42501.
XX
XX New human NOVX polypeptides and nucleic acid molecules, useful for
PT diagnosing, preventing or treating NOVX-associated disorder, e.g. cancer,
PT atherosclerosis, diabetes, Alzheimer's disease, Parkinson's disease or
PT scleroderma.
XX
PS Claim 20; SEQ ID NO 349; 610pp; English.
XX
CC The invention relates to human NOVX polypeptides and the polynucleotides
CC encoding them. The invention also relates to antibodies specific to the
CC NOVX polypeptides. The polypeptides, polynucleotides and antibodies are
CC useful for manufacturing a medicament for treating a syndrome associated
CC with a human disease, such as a pathology associated with the NOVX
CC polypeptide. The sequences are useful for diagnosing, treating or
CC preventing a NOVX-associated disorder, e.g., cancer, atherosclerosis,
CC diabetes, Alzheimer's disease, Parkinson's disease, graft-versus-host
CC disease, scleroderma, hypertension, haemophilia, idiopathic
CC thrombocytopenic purpura, immunodeficiencies, AIDS, dyslipidemia,
CC obesity, Crohn's disease, bronchial asthma, anorexia, cancer-associated
CC cachexia, multiple sclerosis or fertility. The nucleic acids may be used
CC as hybridisation probes, in chromosome mapping, in tissue typing, in
CC preventive medicine or in pharmacogenomics. This sequence represents a
CC human NOVX polynucleotide of the invention.
XX
SQ Sequence 2017 BP; 345 A; 613 C; 617 G; 442 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 12; Length 2017;
Best Local Similarity 100.0%; Pred. No. 2.4e-24;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AAGCGCGCTCGAGTCTAGAGGCGCGCTTTAAACCGCTGATCAGCTCGACTGTGCT 60
Db 1839 AAGCGCGCGCTCGAGTCTAGAGGCGCGCTTTAAACCGCTGATCAGCTCGACTGTGCT 1898
Qy 61 TCTAGTTCAGCCATCTGTTGTTGCTCCCTCCCGTGC 101
|||||
```

Db 1899 TCTAGTTGCCAGCATCTGTTGTTGGCCCTCCCGGCGCC 1939

RESULT 5
ADO42506

XX ADO42506 standard; cDNA; 2022 BP.
XX AC ADO42506;
XX DT 15-JUL-2004 (first entry)
XX DE Human NOVX polynucleotide #178.
XX KW Human; NOVX; gene; ss; cancer; atherosclerosis; diabetes;
KW Alzheimer's disease; Parkinson's disease; graft-versus-host disease;
KW scleroderma; hypertension; haemophilia;
KW idiopathic thrombocytopenic purpura; immunodeficiency; AIDS;
KW dyslipidemia; obesity; Crohn's disease; bronchial asthma; anorexia;
KW cancer-associated cachexia; multiple sclerosis; fertility.
XX OS Homo sapiens.
XX PN US2004058338-A1.
XX PD 25-MAR-2004.
XX PF 02-DEC-2002; 2002US-00307817.
XX PR 03-DEC-2001; 2001US-0336881P.
PR 05-DEC-2001; 2001US-0336820P.
PR 07-DEC-2001; 2001US-0338285P.
PR 07-DEC-2001; 2001US-0338338P.
PR 10-DEC-2001; 2001US-0338989P.
PR 10-DEC-2001; 2001US-0339022P.
PR 11-DEC-2001; 2001US-0339314P.
PR 11-DEC-2001; 2001US-0339516P.
PR 11-DEC-2001; 2001US-0339517P.
PR 11-DEC-2001; 2001US-0339611P.
PR 12-DEC-2001; 2001US-0340981P.
PR 12-DEC-2001; 2001US-0341346P.
PR 14-DEC-2001; 2001US-0340390P.
PR 14-DEC-2001; 2001US-0340440P.
PR 14-DEC-2001; 2001US-0340565P.
PR 14-DEC-2001; 2001US-0340608P.
PR 14-DEC-2001; 2001US-0341144P.
PR 17-DEC-2001; 2001US-0341477P.
PR 17-DEC-2001; 2001US-0341540P.
PR 18-DEC-2001; 2001US-0341768P.
PR 20-DEC-2001; 2001US-0342592P.
PR 31-DEC-2001; 2001US-0344903P.
PR 01-FEB-2002; 2002US-0353286P.
PR 01-FEB-2002; 2002US-0353288P.
PR 26-FEB-2002; 2002US-0359599P.
PR 26-FEB-2002; 2002US-0359626P.
PR 26-FEB-2002; 2002US-0359671P.
PR 27-FEB-2002; 2002US-0359914P.
PR 27-FEB-2002; 2002US-0359956P.
PR 28-FEB-2002; 2002US-0360924P.
PR 28-FEB-2002; 2002US-0360964P.
PR 28-FEB-2002; 2002US-0361028P.
PR 28-FEB-2002; 2002US-0361256P.
PR 28-FEB-2002; 2002US-0361264P.
PR 05-MAR-2002; 2002US-0361770P.
PR 05-MAR-2002; 2002US-0362230P.
PR 13-MAR-2002; 2002US-0364181P.
PR 13-MAR-2002; 2002US-0364238P.
PR 15-MAR-2002; 2002US-0364978P.
PR 15-MAR-2002; 2002US-0365025P.
PR 17-APR-2002; 2002US-0373288P.
PR 15-MAY-2002; 2002US-0380981P.
PR 16-MAY-2002; 2002US-0381004P.
PR 17-MAY-2002; 2002US-0381495P.
PR 28-MAY-2002; 2002US-0383534P.

PR 28-MAY-2002; 2002US-0383744P.
PR 29-MAY-2002; 2002US-0383829P.
PR 29-MAY-2002; 2002US-0384024P.
PR 02-JUL-2002; 2002US-0393332P.
PR 06-AUG-2002; 2002US-0401315P.
PR 07-AUG-2002; 2002US-0401788P.
PR 20-AUG-2002; 2002US-0404676P.
PR 23-AUG-2002; 2002US-0405400P.
PR 23-AUG-2002; 2002US-0405684P.
PR 23-AUG-2002; 2002US-0405687P.
PR 23-AUG-2002; 2002US-0405698P.
PR 26-AUG-2002; 2002US-0406353P.
XX
PA (AGEE/) AGEE M L.
PA (ALSO/) ALSOBROOK J P.
PA (ANDE/) ANDERSON D W.
PA (BERG/) BERGHS C.
PA (BOLD/) BOLDOG F L.
PA (BURG/) BURGESS C B.
PA (CATT/) CATTERTON E.
PA (DIPI/) DIPIPPO V A.
PA (EDIN/) EDINGER S R.
PA (EISE/) EISEN A.
PA (ELLE/) ELLERMAN K.
PA (GANG/) GANGOLLI E A.
PA (GERL/) GERLACH V.
PA (GORM/) GORMAN L.
PA (ROTH/) ROTHBERG B G.
PA (GUOX/) GUO X S.
PA (HERR/) HERRMANN J L.
PA (HALV/) HALVORSEN Y.
PA (JIWW/) JI W.
PA (KEKU/) KEKUDA R.
PA (KHRA/) KHRAMTSOV N V.
PA (LARO/) LAROCHELLE W J.
PA (LEPL/) LEPLEY D M.
PA (LILL/) LI L.
PA (MACD/) MACDOUGALL J R.
PA (MILL/) MILLER C E.
PA (ORTT/) ORT T.
PA (PADI/) PADIGARU M.
PA (PATT/) PATTURAJAN M.
PA (PENA/) PENNA C E A.
PA (PEYM/) PEYMAN J A.
PA (RIEG/) RIEGER D K.
PA (ROTH/) ROTHENBERG M E.
PA (SHEN/) SHENOY S G.
PA (SMIT/) SMITHSON G.
PA (SPAD/) SPADERNA S K.
PA (SPYT/) SPYTEK K A.
PA (STON/) STONE D J.
PA (TAUP/) TAUPIER R J.
PA (VERN/) VERNET C A M.
PA (VOSS/) VOSS E Z.
PA (ZHON/) ZHONG M.
XX
PI Agee ML, Alsobrook JP, Anderson DW, Berghs C, Boldog FL;
PI Burgess CE, Catterton E, Dipippo VA, Edinger SR, Eisen A;
PI Ellerman K, Gangolli EA, Gerlach V, Gorman L, Rothberg BG, Guo XS;
PI Hermann JL, Halvorsen Y, Ji W, Kekuda R, Khrantsov NV;
PI Larochele WJ, Lepley DM, Li L, Macdougall JR, Miller CE, Ort T;
PI Padigar M, Patturajan M, Pena CE, Peyman JA, Rieger DK;
PI Rothenberg ME, Shenoy SG, Smithson G, Spaderna SK, Spytek KA;
PI Stone DJ, Taupier RJ, Vernet CAM, Voss EZ, Zhong M;
XX
XX WPI: 2004-268786/25.
XX P-PSDB; ADO42507.
XX
XX New human NOVX polypeptides and nucleic acid molecules, useful for
XX diagnosing, preventing or treating NOVX-associated disorder, e.g. cancer,
XX atherosclerosis, diabetes, Alzheimer's disease, Parkinson's disease or
XX scleroderma.

```
PS Claim 20; SEQ ID NO 355; 610pp; English.
XX
XX The invention relates to human NOVX polypeptides and the polynucleotides
CC encoding them. The invention also relates to antibodies specific to the
CC NOVX polypeptides. The polypeptides, polynucleotides and antibodies are
CC useful for manufacturing a medicament for treating a syndrome associated
CC with a human disease, such as a pathology associated with the NOVX
CC polypeptide. The sequences are useful for diagnosing, treating or
CC preventing a NOVX-associated disorder, e.g., cancer, atherosclerosis,
CC diabetes, Alzheimer's disease, Parkinson's disease, graft-versus-host
CC disease, scleroderma, hypertension, haemophilia, idiopathic
CC thrombocytopenic purpura, immunodeficiencies, AIDS, dyslipidemia,
CC obesity, Crohn's disease, bronchial asthma, anorexia, cancer-associated
CC cachexia, multiple sclerosis or fertility. The nucleic acids may be used
CC as hybridisation probes, in chromosome mapping, in tissue typing, in
CC preventive medicine or in pharmacogenomics. This sequence represents a
CC human NOVX polynucleotide of the invention.
XX
SQ Sequence 2022 BP; 347 A; 614 C; 618 G; 443 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 12; Length 2022;
Best Local Similarity 100.0%; Pred. No. 2.4e-24;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAGCGCGCGTCGAGCTCTAGAGGGCGCCGTTTAAACCCGCTGATCAGCTCGACTGTGCCT 60
DB 1843 AAGCGCGCGTCGAGCTCTAGAGGGCGCCGTTTAAACCCGCTGATCAGCTCGACTGTGCCT 1902
QY 61 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGCGTGC 101
DB 1903 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGCGTGC 1943
RESULT 6
AAD10237
ID AAD10237 standard; DNA; 6050 BP.
XX
XX AAD10237;
XX
XX 24-SEP-2001 (first entry)
XX
XX Commercial plasmid vector pVAX1/LacZ.
XX
XX Plasmid; cyclic; circular; replicon; exogenous gene; marker gene;
XX transcription termination; immunostimulatory sequence; ISS; antiviral;
XX non-essential nucleotide; molecular biology application; gene therapy;
XX DNA vaccine; cloning; gene expression; in vitro protein production;
XX cytosolic; pVAX1/LacZ; cytomegalovirus promoter; lacZ gene;
XX kanamycin resistant; ds.
XX
XX Cytomegalovirus.
XX Unidentified.
XX Chimeric.
XX
XX Key Location/Qualifiers
XX misc_feature 1..112
XX /tag= a
XX /note= "Corresponds to the non-essential nucleotide
XX sequence that have been removed in the novel DNA plasmid
XX vector"
XX misc_feature 5092..5143
XX /tag= b
XX /note= "3', non-coding region of lacZ gene from pUC18 that
XX corresponds to the non-essential nucleotide sequence that
XX have been removed in the novel DNA plasmid vector"
XX misc_feature 5144..5258
XX /tag= c
XX /note= "Ampicillin promoter that corresponds to the non-
XX essential nucleotide sequence that have been removed in
XX the novel DNA plasmid vector"
XX
XX WO200151626-A2.
XX
```

```
PD 19-JUL-2001.
XX
XX 09-JAN-2001; 2001WO-US001255.
XX
XX 10-JAN-2000; 2000US-00480879.
XX
XX (ELIM-) ELIM BIOPHARMACEUTICALS INC.
XX
XX Lu X, Sun L, Zhang Y;
XX
XX WPI; 2001-451855/48.
XX
XX New plasmid DNA vectors, useful for most molecular biology applications,
XX e.g. gene therapy, DNA vaccines, cloning and expression of genes, and in
XX the in vitro production of polypeptides and/or proteins.
XX
XX Example 2; Page 39-41; 50pp; English.
XX
XX The present invention relates to plasmid DNA vectors comprising
XX essentially of a replicon and at least one other component selected from
XX promoter, intron, exogenous gene, transcription termination sequence,
XX selectable marker gene, detectable marker gene and an immunostimulatory
XX sequence (ISS), where the non-essential nucleotide sequences have been
XX substantially removed from these vectors. The plasmid DNA vectors are
XX useful in most molecular biology applications, e.g. gene therapy, DNA
XX vaccines, cloning and expression of genes, and in the in vitro production
XX of polypeptides and/or proteins. The present sequence is a commercial
XX plasmid DNA vector pVAX1/LacZ which comprises Cytomegalovirus promoter,
XX the lacZ gene, pUC origin of replication and the kanamycin resistance
XX gene
XX
XX Sequence 6050 BP; 1346 A; 1597 C; 1696 G; 1411 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 5; Length 6050;
Best Local Similarity 100.0%; Pred. No. 3.2e-24;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAGCGCGCGTCGAGCTCTAGAGGGCGCCGTTTAAACCCGCTGATCAGCTCGACTGTGCCT 60
DB 3828 AAGCGCGCGTCGAGCTCTAGAGGGCGCCGTTTAAACCCGCTGATCAGCTCGACTGTGCCT 3887
QY 61 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGCGTGC 101
DB 3888 TCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCGCGTGC 3928
RESULT 7
AAD06054
ID AAD06054 standard; DNA; 6567 BP.
XX
XX AAD06054;
XX
XX 31-JUL-2001 (first entry)
XX
XX Plasmid E2CLEDAS encoding fusion protein comprising E2C ZFP-ER LBD-TA.
XX
XX Plasmid E2CLEDAS; fusion protein; nucleotide-binding domain; NBD;
XX ligand-binding domain; LBD; transcription regulating domain; TRD;
XX zinc finger protein; ZFP; ligand-activated transcriptional regulator;
XX gene regulation; gene therapy; cell proliferative disorder; cancer;
XX psoriasis; pemphigus vulgaris; Behcet's syndrome; lipid histiocytosis;
XX E2C ZFP; human; oestrogen receptor; ER; VP16; TA; transactivation domain;
XX cyclic; circular; ds.
XX
XX Unidentified.
XX Homo sapiens.
XX OS
XX Herpes simplex virus.
XX OS
XX Cytomegalovirus.
XX OS
XX Enterobacteria phage T7.
XX OS
XX Rhesus macaque polyoma virus.
XX OS
XX Chimeric.
XX
XX WO200130843-A1.
XX
```

```

XX 03-MAY-2001.
PD
XX
XX
XX 23-OCT-2000; 2000WO-EP010430.
XX
XX 25-OCT-1999; 99US-00433042.
XX
XX 02-JUN-2000; 2000US-00586625.
XX
XX (NOVS ) NOVARTIS AG.
XX (SCRI ) SCRIPPS RES INST.
XX
XX Barbas CF, Kadan M, Beerli R;
XX WPI; 2001-308618/32.
XX
XX New fusion protein containing nucleotide-binding and ligand-binding
XX domains, useful e.g. in gene therapy of cancer, provides ligand-activated
XX control of gene expression.
XX
XX Claim 25; Page 186-187; 218pp; English.
XX
XX The invention relates to fusion protein comprising a nucleotide-binding
XX domain (NBD), a ligand-binding domain (LBD) of an intracellular receptor
XX (ICR) and a transcription regulating domain (TRD). NBD is a polydactyl
XX zinc finger protein (ZFP), or a modular part of it, that interacts
XX specifically with a contiguous sequence of at least 3 nucleotides. The
XX fusion protein functions as a ligand-activated transcriptional regulator.
XX The fusion protein and the nucleic acid encoding it, are used to regulate
XX gene expression, particularly in gene therapy for treating malignant cell
XX proliferative diseases (e.g. colon cancer, prostate cancer, renal-cell
XX carcinoma) and non-malignant cell proliferative diseases (e.g. psoriasis,
XX pemphigus vulgaris, Behcet's syndrome and lipid histiocytosis). The
XX fusion protein and its DNA are also useful for treating diseases caused
XX by viruses in humans/plants, genetic and/or acquired diseases. The fusion
XX protein can be designed to target any selected gene (endogenous or
XX exogenous), and can be made to have different selectivity or specificity
XX for endogenous or exogenous ligands. The present sequence is E2CLBDAS
XX construct encoding fusion protein comprising E2C zinc finger protein
XX receptor LBD fragment and VP16 transactivation domain (TA) from Herpes
XX simplex virus. The ZFP serves as NBD and VP16 TA domain functions as
XX transcription activator. The E2CLBDAS construct is based on plasmid
XX pCDNA3.1 that comprises sequences from cytomegalovirus, bacteriophage T7
XX and simian virus 40 (SV40)
XX
XX Sequence 6567 BP; 1523 A; 1742 C; 1683 G; 1619 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 101; DB 4; Length 6567;
XX Best Local Similarity 100.0%; Pred. No. 3.2e-24;
XX Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 AAGCGCGCGTTCGAGTCTAGAGGCGCGGTTTAAACCGGTGATCAGCTCGACTGTGCCT 60
XX |||||
XX 2171 AAGCGCGCGTTCGAGTCTAGAGGCGCGGTTTAAACCGGTGATCAGCTCGACTGTGCCT 2230
XX
XX 61 TCTAGTTGCCAGCCATCTGTGTTTGGCCCTCCCGGTGCC 101
XX |||||
XX 2231 TCTAGTTGCCAGCCATCTGTGTTTGGCCCTCCCGGTGCC 2271
XX
XX RESULT 8
XX AAD06049
XX ID AAD06049 standard; DNA; 6623 BP.
XX
XX AC AAD06049;
XX
XX 31-JUL-2001 (first entry)
XX
XX Plasmid C7LBDAS encoding fusion protein comprising C7 ZFP-ER LBD-TA.
XX
XX Plasmid C7LBDAS; fusion protein; nucleotide-binding domain; NBD;
XX ligand-binding domain; LBD; transcription regulating domain; TRD;
XX zinc finger protein; ZFP; ligand-activated transcriptional regulator;

```

```

KW gene regulation; gene therapy; cell proliferative disorder; cancer;
KW psoriasis; pemphigus vulgaris; Behcet's syndrome; lipid histiocytosis;
KW C7 ZFP; Zif268; murine; human; oestrogen receptor; ER; VP16; TA;
KW transactivation domain; cyclic; circular; ds.
XX
XX Mus sp.
XX
XX Homo sapiens.
XX
XX Herpes simplex virus.
XX
XX Cytomegalovirus.
XX
XX Enterobacteria phage T7.
XX
XX Rhesus macaque polyoma virus.
XX
XX Unidentified.
XX
XX Chimeric.
XX
XX WO200130843-A1.
XX
XX 03-MAY-2001.
XX
XX 23-OCT-2000; 2000WO-EP010430.
XX
XX 25-OCT-1999; 99US-00433042.
XX
XX 02-JUN-2000; 2000US-00586625.
XX
XX (NOVS ) NOVARTIS AG.
XX (SCRI ) SCRIPPS RES INST.
XX
XX Barbas CF, Kadan M, Beerli R;
XX WPI; 2001-308618/32.
XX
XX New fusion protein containing nucleotide-binding and ligand-binding
XX domains, useful e.g. in gene therapy of cancer, provides ligand-activated
XX control of gene expression.
XX
XX Claim 25; Page 176-178; 218pp; English.
XX
XX The invention relates to fusion protein comprising a nucleotide-binding
XX domain (NBD), a ligand-binding domain (LBD) of an intracellular receptor
XX (ICR) and a transcription regulating domain (TRD). NBD is a polydactyl
XX zinc finger protein (ZFP), or a modular part of it, that interacts
XX specifically with a contiguous sequence of at least 3 nucleotides. The
XX fusion protein functions as a ligand-activated transcriptional regulator.
XX The fusion protein and the nucleic acid encoding it, are used to regulate
XX gene expression, particularly in gene therapy for treating malignant cell
XX proliferative diseases (e.g. colon cancer, prostate cancer, renal-cell
XX carcinoma) and non-malignant cell proliferative diseases (e.g. psoriasis,
XX pemphigus vulgaris, Behcet's syndrome and lipid histiocytosis). The
XX fusion protein and its DNA are also useful for treating diseases caused
XX by viruses in humans/plants, genetic and/or acquired diseases. The fusion
XX protein can be designed to target any selected gene (endogenous or
XX exogenous), and can be made to have different selectivity or specificity
XX for endogenous or exogenous ligands. The present sequence is C7LBDAS
XX construct encoding fusion protein comprising C7 zinc finger protein (ZFP)
XX which is a variant of murine Cys2-His2 ZFP Zif268, human oestrogen
XX receptor LBD fragment and VP16 transactivation domain (TA) from Herpes
XX simplex virus. The ZFP serves as NBD and VP16 TA domain functions as
XX transcription activator. The C7LBDAS construct is based on plasmid
XX pCDNA3.1 that comprises sequences from cytomegalovirus, bacteriophage T7
XX and simian virus 40 (SV40)
XX
XX Sequence 6623 BP; 1530 A; 1754 C; 1703 G; 1636 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 101; DB 4; Length 6623;
XX Best Local Similarity 100.0%; Pred. No. 3.2e-24;
XX Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 AAGCGCGCGTTCGAGTCTAGAGGCGCGGTTTAAACCGGTGATCAGCTCGACTGTGCCT 60
XX |||||
XX 2227 AAGCGCGCGTTCGAGTCTAGAGGCGCGGTTTAAACCGGTGATCAGCTCGACTGTGCCT 2286
XX
XX 61 TCTAGTTGCCAGCCATCTGTGTTTGGCCCTCCCGGTGCC 101
XX |||||
XX 2287 TCTAGTTGCCAGCCATCTGTGTTTGGCCCTCCCGGTGCC 2327
XX

```

RESULT 9
 AAD06055
 ID AAD06055 standard; DNA; 6639 BP.
 AC
 AC AAD06055;
 XX
 DT 31-JUL-2001 (first entry)
 XX
 DE Plasmid E2CLBDBS encoding fusion protein comprising E2C ZFP-ER LBD-TA.
 XX
 KW Plasmid E2CLBDBS; fusion protein; nucleotide-binding domain; NBD;
 KW ligand-binding domain; LBD; transcription regulating domain; TRD;
 KW zinc finger protein; ZFP; ligand-activated transcriptional regulator;
 KW gene regulation; gene therapy; cell proliferative disorder; cancer;
 KW psoriasis; pemphigus vulgaris; Behcet's syndrome; lipid histiocytosis;
 KW E2C ZFP; human; oestrogen receptor; ER; VP16; TA; transactivation domain;
 KW cyclic; circular; ds.
 XX
 OS Unidentified.
 OS Homo sapiens.
 OS Herpes simplex virus.
 OS Cytomegalovirus.
 OS Enterobacteria phage T7.
 OS Rhesus macaque polyoma virus.
 OS Chimeric.
 XX
 PN WO200130843-A1.
 XX
 XX
 PD 03-MAY-2001.
 XX
 PF 23-OCT-2000; 2000WO-EP010430.
 XX
 PR 25-OCT-1999; 99US-00433042.
 PR 02-JUN-2000; 2000US-00586625.
 XX
 PA (NOVS) NOVARTIS AG.
 PA (SCRI) SCRIPPS RES INST.
 XX
 PI Barbas CF, Kadan M, Beerli R;
 XX WPI; 2001-308618/32.
 DR
 XX
 PT New fusion protein containing nucleotide-binding and ligand-binding
 PT domains, useful e.g. in gene therapy of cancer, provides ligand-activated
 PT control of gene expression.
 XX
 PS Claim 25; Page 188-189; 218pp; English.
 XX
 CC The invention relates to fusion protein comprising a nucleotide-binding
 CC domain (NBD), a ligand-binding domain (LBD) of an intracellular receptor
 CC (ICR) and a transcription regulating domain (TRD). NBD is a polydactyl
 CC zinc finger protein (ZFP), or a modular part of it, that interacts
 CC specifically with a contiguous sequence of at least 3 nucleotides. The
 CC fusion protein functions as a ligand-activated transcriptional regulator.
 CC The fusion protein and the nucleic acid encoding it, are used to regulate
 CC gene expression, particularly in gene therapy for treating malignant cell
 CC proliferative diseases (e.g. colon cancer, prostate cancer, renal-cell
 CC carcinoma) and non-malignant cell proliferative diseases (e.g. psoriasis,
 CC pemphigus vulgaris, Behcet's syndrome and lipid histiocytosis). The
 CC fusion protein and its DNA are also useful for treating diseases caused
 CC by viruses in humans/plants, genetic and/or acquired diseases. The fusion
 CC protein can be designed to target any selected gene (endogenous or
 CC exogenous), and can be made to have different selectivity or specificity
 CC for endogenous or exogenous ligands. The present sequence is E2CLBDBS
 CC construct encoding fusion protein comprising E2C zinc finger protein
 CC (ZFP) that binds human erbB-2 target sequence e2c, human oestrogen
 CC receptor LBD fragment and VP16 transactivation domain (TA) from Herpes
 CC simplex virus. The ZFP serves as NBD and VP16 TA domain functions as
 CC transcription activator. The E2CLBDBS construct is based on plasmid
 CC pcDNA3.1 that comprises sequences from cytomegalovirus, bacteriophage T7
 CC and simian virus 40 (SV40)

XX
 SQ Sequence 6639 BP; 1546 A; 1749 C; 1718 G; 1626 T; 0 U; 0 Other;
 Query Match 100.0%; Score 101; DB 4; Length 6639;
 Best Local Similarity 100.0%; Pred. No. 3.2e-24;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AAGCGGCCGCTCGAGTCTAGAGGGCCCGTTAAACCCGCTGATCAGCTCGACTGTGCT 60
 DB 2243 AAGCGGCCGCTCGAGTCTAGAGGGCCCGTTAAACCCGCTGATCAGCTCGACTGTGCT 2302
 QY 61 TCTAGTTGCCAGCATCTGTTGTTGCCCTCCCGCGCC 101
 DB 2303 TCTAGTTGCCAGCATCTGTTGTTGCCCTCCCGCGCC 2343
 RESULT 10
 AAD06051
 ID AAD06051 standard; DNA; 6695 BP.
 XX
 AC AAD06051;
 XX
 DT 31-JUL-2001 (first entry)
 XX
 DE Plasmid C7LBDBS encoding fusion protein comprising C7 ZFP-ER LBD-TA.
 XX
 KW Plasmid C7LBDBS; fusion protein; nucleotide-binding domain; NBD;
 KW ligand-binding domain; LBD; transcription regulating domain; TRD;
 KW zinc finger protein; ZFP; ligand-activated transcriptional regulator;
 KW gene regulation; gene therapy; cell proliferative disorder; cancer;
 KW psoriasis; pemphigus vulgaris; Behcet's syndrome; lipid histiocytosis;
 KW C7 ZFP; Zif268; murine; human; oestrogen receptor; ER; VP16; TA;
 KW transactivation domain; cyclic; circular; ds.
 XX
 OS Mus sp.
 OS Homo sapiens.
 OS Herpes simplex virus.
 OS Cytomegalovirus.
 OS Enterobacteria phage T7.
 OS Rhesus macaque polyoma virus.
 OS Unidentified.
 OS Chimeric.
 XX
 PN WO200130843-A1.
 XX
 XX
 PD 03-MAY-2001.
 XX
 PF 23-OCT-2000; 2000WO-EP010430.
 XX
 PR 25-OCT-1999; 99US-00433042.
 PR 02-JUN-2000; 2000US-00586625.
 XX
 PA (NOVS) NOVARTIS AG.
 PA (SCRI) SCRIPPS RES INST.
 XX
 PI Barbas CF, Kadan M, Beerli R;
 XX WPI; 2001-308618/32.
 DR
 XX
 PT New fusion protein containing nucleotide-binding and ligand-binding
 PT domains, useful e.g. in gene therapy of cancer, provides ligand-activated
 PT control of gene expression.
 XX
 PS Claim 25; Page 180-182; 218pp; English.
 XX
 CC The invention relates to fusion protein comprising a nucleotide-binding
 CC domain (NBD), a ligand-binding domain (LBD) of an intracellular receptor
 CC (ICR) and a transcription regulating domain (TRD). NBD is a polydactyl
 CC zinc finger protein (ZFP), or a modular part of it, that interacts
 CC specifically with a contiguous sequence of at least 3 nucleotides. The
 CC fusion protein functions as a ligand-activated transcriptional regulator.
 CC The fusion protein and the nucleic acid encoding it, are used to regulate
 CC gene expression, particularly in gene therapy for treating malignant cell

CC proliferative diseases (e.g. colon cancer, prostate cancer, renal-cell
CC carcinoma) and non-malignant cell proliferative diseases (e.g. psoriasis,
CC pemphigus vulgaris, Behcet's syndrome and lipid histiocytosis). The
CC fusion protein and its DNA are also useful for treating diseases caused
CC by viruses in humans/plants, genetic and/or acquired diseases. The fusion
CC protein can be designed to target any selected gene (endogenous or
CC exogenous), and can be made to have different selectivity or specificity
CC for endogenous or exogenous ligands. The present sequence is C7LDBS
CC construct encoding fusion protein comprising C7 zinc finger protein (ZFP)
CC which is a variant of murine Cys2-His2 ZFP Zif268, human oestrogen
CC receptor LBD fragment and VP16 transactivation domain (TA) from Herpes
CC simplex virus. The ZFP serves as NBD and VP16 TA domain functions as
CC transcription activator. The C7LDBS construct is based on plasmid
CC pCDNA3.1 that comprises sequences from cytomegalovirus, bacteriophage T7
CC and simian virus 40 (SV40)

XX SQ Sequence 6695 BP; 1553 A; 1762 C; 1737 G; 1643 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 4; Length 6695;
Best Local Similarity 100.0%; Pred. No. 3.2e-24;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGCGCGCTCGAGTCTAGAGGGCCCGTTTAAACCCGCTGATCAGCTCGACTGTGCCT 60
Db 2299 AAGCGCGCTCGAGTCTAGAGGGCCCGTTTAAACCCGCTGATCAGCTCGACTGTGCCT 2358

Qy 61 TCTAGTTGCCAGCATCTGTTGTTGCCCTCCCGTGCCT 101
Db 2359 TCTAGTTGCCAGCATCTGTTGTTGCCCTCCCGTGCCT 2399

RESULT 11
AAD06057
ID AAD06057 standard; DNA; 6695 BP.
XX AC AAD06057;
XX DT 31-JUL-2001 (first entry)
XX DE Plasmid C7LDBSG400V encoding fusion protein.
XX KW Plasmid C7LDBSG400V; fusion protein; nucleotide-binding domain; NBD;
KW ligand-binding domain; LBD; transcription regulating domain; TRD;
KW zinc finger protein; ZFP; ligand-activated transcriptional regulator;
KW gene regulation; gene therapy; cell proliferative disorder; cancer;
KW psoriasis; pemphigus vulgaris; Behcet's syndrome; lipid histiocytosis;
KW C7 ZFP; Zif268; murine; human; oestrogen receptor; ER; mutant; mutein;
KW VP16; TA; transactivation domain; cyclic; circular; ds.
XX OS Mus sp.
OS Homo sapiens.
OS Synthetic.
OS Herpes simplex virus.
OS Cytomegalovirus.
OS Enterobacteria phage T7.
OS Rhesus macaque polyoma virus.
XX PN WO200130843-A1.
XX XX
XX FD 03-MAY-2001.
XX PF 23-OCT-2000; 2000WO-EP010430.
XX PR 25-OCT-1999; 98US-00433042.
XX PR 02-JUN-2000; 2000US-00586625.
XX XX
XX FA (NOVS) NOVARTIS AG.
XX FA (SCRI) SCRIPPS RES INST.
XX PI Barbas CF, Kadan M, Beerli R;

DR WPI; 2001-308618/32.
XX New fusion protein containing nucleotide-binding and ligand-binding
PT domains, useful e.g. in gene therapy of cancer, provides ligand-activated
PT control of gene expression.
XX Claim 25; Page 191-193; 218pp; English.
XX The invention relates to fusion protein comprising a nucleotide-binding
CC domain (NBD), a ligand-binding domain (LBD) of an intracellular receptor
CC (ICR) and a transcription regulating domain (TRD). NBD is a polydactyl
CC zinc finger protein (ZFP), or a modular part of it, that interacts
CC specifically with a contiguous sequence of at least 3 nucleotides. The
CC fusion protein functions as a ligand-activated transcriptional regulator.
CC The fusion protein and the nucleic acid encoding it, are used to regulate
CC gene expression, particularly in gene therapy for treating malignant cell
CC proliferative diseases (e.g. colon cancer, prostate cancer, renal-cell
CC carcinoma) and non-malignant cell proliferative diseases (e.g. psoriasis,
CC pemphigus vulgaris, Behcet's syndrome and lipid histiocytosis). The
CC fusion protein and its DNA are also useful for treating diseases caused
CC by viruses in humans/plants, genetic and/or acquired diseases. The fusion
CC protein can be designed to target any selected gene (endogenous or
CC exogenous), and can be made to have different selectivity or specificity
CC for endogenous or exogenous ligands. The present sequence is C7LDBSG400V
CC construct encoding fusion protein comprising C7 zinc finger protein (ZFP)
CC which is a variant of murine Cys2-His2 ZFP Zif268, human oestrogen
CC receptor LBD fragment containing G400V mutation and Herpes simplex virus
CC VP16 transactivation domain (TA) The ZFP serves as NBD and VP16 TA domain
CC functions as transcription activator. The C7LDBSG400V construct is based
CC on plasmid pCDNA3.1 that comprises sequences from cytomegalovirus,
CC bacteriophage T7 and simian virus 40 (SV40)

XX SQ Sequence 6695 BP; 1553 A; 1762 C; 1736 G; 1644 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 4; Length 6695;
Best Local Similarity 100.0%; Pred. No. 3.2e-24;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAGCGCGCTCGAGTCTAGAGGGCCCGTTTAAACCCGCTGATCAGCTCGACTGTGCCT 60
Db 2299 AAGCGCGCTCGAGTCTAGAGGGCCCGTTTAAACCCGCTGATCAGCTCGACTGTGCCT 2358

Qy 61 TCTAGTTGCCAGCATCTGTTGTTGCCCTCCCGTGCCT 101
Db 2359 TCTAGTTGCCAGCATCTGTTGTTGCCCTCCCGTGCCT 2399

RESULT 12
AAD06058
ID AAD06058 standard; DNA; 6695 BP.
XX AC AAD06058;
XX DT 31-JUL-2001 (first entry)
XX DE Plasmid C7LDBSG521R encoding fusion protein.
XX KW Plasmid C7LDBSG521R; fusion protein; nucleotide-binding domain; NBD;
KW ligand-binding domain; LBD; transcription regulating domain; TRD;
KW zinc finger protein; ZFP; ligand-activated transcriptional regulator;
KW gene regulation; gene therapy; cell proliferative disorder; cancer;
KW psoriasis; pemphigus vulgaris; Behcet's syndrome; lipid histiocytosis;
KW C7 ZFP; Zif268; murine; human; oestrogen receptor; ER; mutant; mutein;
KW VP16; TA; transactivation domain; cyclic; circular; ds.
XX OS Mus sp.
OS Homo sapiens.
OS Synthetic.
OS Herpes simplex virus.
OS Cytomegalovirus.
OS Enterobacteria phage T7.
OS Rhesus macaque polyoma virus.
XX OS Unidentified.

OS Chimeric.
FN WO200130843-A1.
XX
XX
PD 03-MAY-2001.
XX
XX 23-OCT-2000; 2000WO-EP010430.
XX
XX 25-OCT-1999; 99US-00433042.
PR 02-JUN-2000; 2000US-00586625.
XX
XX (NOVS) NOVARTIS AG.
PA (SCRI) SCRIPPS RES INST.
XX
XX Barbas CF, Kadan M, Beerli R;
XX WPI; 2001-308618/32.
XX
XX New fusion protein containing nucleotide-binding and ligand-binding
PT domains, useful e.g. in gene therapy of cancer, provides ligand-activated
PT control of gene expression.
XX
XX Claim 25; Page 193-195; 218pp; English.
XX
XX The invention relates to fusion protein comprising a nucleotide-binding
CC domain (NBD), a ligand-binding domain (LBD) of an intracellular receptor
CC (ICR) and a transcription regulating domain (TRD). NBD is a polydactyl
CC zinc finger protein (ZFP), or a modular part of it, that interacts
CC specifically with a contiguous sequence of at least 3 nucleotides. The
CC fusion protein functions as a ligand-activated transcriptional regulator.
CC The fusion protein and the nucleic acid encoding it, are used to regulate
CC gene expression, particularly in gene therapy for treating malignant cell
CC proliferative diseases (e.g. colon cancer, prostate cancer, renal-cell
CC carcinoma) and non-malignant cell proliferative diseases (e.g. psoriasis,
CC pemphigus vulgaris, Behcet's syndrome and lipid histiocytosis). The
CC fusion protein and its DNA are also useful for treating diseases caused
CC by viruses in humans/plants, genetic and/or acquired diseases. The fusion
CC protein can be designed to target any selected gene (endogenous or
CC exogenous), and can be made to have different selectivity or specificity
CC for endogenous or exogenous ligands. The present sequence is C7LBD5G521R
CC construct encoding fusion protein comprising C7 zinc finger protein (ZFP)
CC which is a variant of murine Cys2-His2 ZFP Zif268, human oestrogen
CC receptor LBD fragment containing G521R mutation and Herpes simplex virus
CC VP16 transactivation domain (TA). The ZFP serves as NBD and VP16 TA domain
CC functions as transcription activator. The C7LBD5G521R construct is based
CC on plasmid pCDNA3.1 that comprises sequences from cytomegalovirus,
CC bacteriophage T7 and simian virus 40 (SV40)
XX
XX Sequence 6695 BP; 1553 A; 1763 C; 1736 G; 1643 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 101; DB 4; Length 6695;
Best Local Similarity 100.0%; Pred. No. 3.2e-24;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AAGCGCGCTCGAGTCTAGAGGGCCCGCTTTAAACCCGCTGATCAGCGCTGACTGTGCT 60
Db 2299 AAGCGCGCTCGAGTCTAGAGGGCCCGCTTTAAACCCGCTGATCAGCGCTGACTGTGCT 2358
Qy 61 TCTAGTTCGACGACATCTGTTGTTGCGCCCTCCCGCGGCC 101
Db 2359 TCTAGTTCGACGACATCTGTTGTTGCGCCCTCCCGCGGCC 2399
RESULT 13
AAD06048
ID AAD06048 standard; DNA; 6746 BP.
XX
XX AAD06048;
XX
XX 31-JUL-2001 (first entry)
DT Plasmid C7LBDAL encoding fusion protein comprising C7 ZFP-ER LBD-TA.
XX
XX

KW Plasmid C7LBDAL; fusion protein; nucleotide-binding domain; NBD;
KW ligand-binding domain; LBD; transcription regulating domain; TRD;
KW zinc finger protein; ZFP; ligand-activated transcriptional regulator;
KW gene regulation; gene therapy; cell proliferative disorder; cancer;
KW psoriasis; pemphigus vulgaris; Behcet's syndrome; lipid histiocytosis;
KW C7 ZFP; Zif268; murine; human; oestrogen receptor; ER; VP16; TA;
KW transactivation domain; cyclic; circular; ds.
XX
XX Mus sp.
OS Homo sapiens.
OS Herpes simplex virus.
OS Cytomegalovirus.
OS Enterobacteria phage T7.
OS Rhesus macaque polyoma virus.
OS Unidentified.
OS Chimeric.
XX
XX WO200130843-A1.
XX
XX 03-MAY-2001.
XX
XX 23-OCT-2000; 2000WO-EP010430.
XX
XX 25-OCT-1999; 99US-00433042.
PR 02-JUN-2000; 2000US-00586625.
XX
XX (NOVS) NOVARTIS AG.
PA (SCRI) SCRIPPS RES INST.
XX
XX Barbas CF, Kadan M, Beerli R;
XX WPI; 2001-308618/32.
XX
XX New fusion protein containing nucleotide-binding and ligand-binding
PT domains, useful e.g. in gene therapy of cancer, provides ligand-activated
PT control of gene expression.
XX
XX Claim 25; Page 174-176; 218pp; English.
XX
XX The invention relates to fusion protein comprising a nucleotide-binding
CC domain (NBD), a ligand-binding domain (LBD) of an intracellular receptor
CC (ICR) and a transcription regulating domain (TRD). NBD is a polydactyl
CC zinc finger protein (ZFP), or a modular part of it, that interacts
CC specifically with a contiguous sequence of at least 3 nucleotides. The
CC fusion protein functions as a ligand-activated transcriptional regulator.
CC The fusion protein and the nucleic acid encoding it, are used to regulate
CC gene expression, particularly in gene therapy for treating malignant cell
CC proliferative diseases (e.g. colon cancer, prostate cancer, renal-cell
CC carcinoma) and non-malignant cell proliferative diseases (e.g. psoriasis,
CC pemphigus vulgaris, Behcet's syndrome and lipid histiocytosis). The
CC fusion protein and its DNA are also useful for treating diseases caused
CC by viruses in humans/plants, genetic and/or acquired diseases. The fusion
CC protein can be designed to target any selected gene (endogenous or
CC exogenous), and can be made to have different selectivity or specificity
CC for endogenous or exogenous ligands. The present sequence is C7LBD5G521R
CC construct encoding fusion protein comprising C7 zinc finger protein (ZFP)
CC which is a variant of murine Cys2-His2 ZFP Zif268, human oestrogen
CC receptor LBD fragment containing G521R mutation and Herpes simplex virus
CC VP16 transactivation domain (TA). The ZFP serves as NBD and VP16 TA domain
CC functions as transcription activator. The C7LBD5G521R construct is based
CC on plasmid pCDNA3.1 that comprises sequences from cytomegalovirus,
CC bacteriophage T7 and simian virus 40 (SV40)
XX
XX Sequence 6746 BP; 1557 A; 1788 C; 1741 G; 1660 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 101; DB 4; Length 6746;
Best Local Similarity 100.0%; Pred. No. 3.3e-24;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AAGCGCGCTCGAGTCTAGAGGGCCCGCTTTAAACCCGCTGATCAGCGCTGACTGTGCT 60
Db 2350 AAGCGCGCTCGAGTCTAGAGGGCCCGCTTTAAACCCGCTGATCAGCGCTGACTGTGCT 2409

```
QY      61 TCTAGTTCAGCCATCTGTTGTTGCGCCCTCCCGTGC 101
Db      2410 TCTAGTTCAGCCATCTGTTGTTGCGCCCTCCCGTGC 2450

RESULT 14
AAD06050
ID      AAD06050 standard; DNA; 6818 BP.
XX
AC      AAD06050;
XX
31-JUL-2001 (first entry)
XX
DE      Plasmid C7LBDL encoding fusion protein comprising C7 ZFP-ER LBD-TA.
XX
KW      Plasmid C7LBDL; fusion protein; nucleotide-binding domain; NBD;
KW      ligand-binding domain; LBD; transcription regulating domain; TRD;
KW      zinc finger protein; ZFP; ligand-activated transcriptional regulator;
KW      gene regulation; gene therapy; cell proliferative disorder; cancer;
KW      psoriasis; pemphigus vulgaris; Behcet's syndrome; lipid histiocytosis;
KW      C7 ZFP; Zif268; murine; human; oestrogen receptor; ER; VP16; TA;
KW      transactivation domain; cyclic; circular; ds.
XX
OS      Mus sp.
OS      Homo sapiens.
OS      Herpes simplex virus.
OS      Cyomegalovirus.
OS      Enterobacteria phage T7.
OS      Rhesus macaque polyoma virus.
OS      Unidentified.
OS      Chimeric.
XX
FN      WO200130843-A1.
XX
PD      03-MAY-2001.
XX
PF      23-OCT-2000; 2000WO-EP010430.
XX
PR      25-OCT-1999; 99US-00433042.
XX
PR      02-JUN-2000; 2000US-00586625.
XX
PA      (NOVS ) NOVARTIS AG.
PA      (SCRI ) SCRIPPS RES INST.
XX
PI      Barbas CF, Kadan M, Beerli R;
XX
DR      WPI; 2001-308618/32.
XX
PT      New fusion protein containing nucleotide-binding and ligand-binding
PT      domains, useful e.g. in gene therapy of cancer, provides ligand-activated
PT      control of gene expression.
XX
FS      Claim 25; Page 178-180; 218pp; English.
XX
CC      The invention relates to fusion protein comprising a nucleotide-binding
CC      domain (NBD), a ligand-binding domain (LBD) of an intracellular receptor
CC      (ICR) and a transcription regulating domain (TRD). NBD is a polydactyl
CC      zinc finger protein (ZFP), or a modular part of it, that interacts
CC      specifically with a contiguous sequence of at least 3 nucleotides. The
CC      fusion protein functions as a ligand-activated transcriptional regulator.
CC      The fusion protein and the nucleic acid encoding it, are used to regulate
CC      gene expression, particularly in gene therapy for treating malignant cell
CC      proliferative diseases (e.g. colon cancer, prostate cancer, renal-cell
CC      carcinoma) and non-malignant cell proliferative diseases (e.g. psoriasis,
CC      pemphigus vulgaris, Behcet's syndrome and lipid histiocytosis). The
CC      fusion protein and its DNA are also useful for treating diseases caused
CC      by viruses in humans/plants, genetic and/or acquired diseases. The fusion
CC      protein can be designed to target any selected gene (endogenous or
CC      exogenous), and can be made to have different selectivity or specificity
CC      for endogenous or exogenous ligands. The present sequence is C7LBDL
CC      construct encoding fusion protein comprising C7 zinc finger protein (ZFP)
CC      which is a variant of murine Cys2-His2 ZFP Zif268, human oestrogen
CC      receptor LBD fragment and VP16 transactivation domain (TA) from Herpes
```

```
CC      simplex virus. The ZFP serves as NBD and VP16 TA domain functions as
CC      transcription activator. The C7LBDL construct is based on plasmid
CC      pCDNA3.1 that comprises sequences from cytomegalovirus, bacteriophage T7
CC      and simian virus 40 (SV40)
XX
SQ      Sequence 6818 BP; 1580 A; 1796 C; 1775 G; 1667 T; 0 U; 0 Other;
Query Match      100.0%; Score 101; DB 4; Length 6818;
Best Local Similarity 100.0%; Pred. No. 3.3e-24;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 AAGCGCGCGCTCGAGTCTAGAGGCGCGGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 60
Db      2422 AAGCGCGCGCTCGAGTCTAGAGGCGCGGTTTAAACCGCTGATCAGCCTCGACTGTGCCT 2481
QY      61 TCTAGTTCAGCCATCTGTTGTTGCGCCCTCCCGTGC 101
Db      2482 TCTAGTTCAGCCATCTGTTGTTGCGCCCTCCCGTGC 2522

RESULT 15
AAD06044
ID      AAD06044 standard; DNA; 6828 BP.
XX
AC      AAD06044;
XX
31-JUL-2001 (first entry)
XX
DE      Plasmid 2C7LBDAS encoding fusion protein comprising 2C7 ZFP-ER LBD-TA.
XX
KW      Plasmid 2C7LBDAS; fusion protein; nucleotide-binding domain; NBD;
KW      ligand-binding domain; LBD; transcription regulating domain; TRD;
KW      zinc finger protein; ZFP; ligand-activated transcriptional regulator;
KW      gene regulation; gene therapy; cell proliferative disorder; cancer;
KW      psoriasis; pemphigus vulgaris; Behcet's syndrome; lipid histiocytosis;
KW      2C7 ZFP; Zif268; murine; human; oestrogen receptor; ER; VP16; TA;
KW      transactivation domain; cyclic; circular; ds.
XX
OS      Mus sp.
OS      Homo sapiens.
OS      Herpes simplex virus.
OS      Cyomegalovirus.
OS      Enterobacteria phage T7.
OS      Rhesus macaque polyoma virus.
OS      Unidentified.
OS      Chimeric.
XX
FN      WO200130843-A1.
XX
PD      03-MAY-2001.
XX
PF      23-OCT-2000; 2000WO-EP010430.
XX
PR      25-OCT-1999; 99US-00433042.
XX
PR      02-JUN-2000; 2000US-00586625.
XX
PA      (NOVS ) NOVARTIS AG.
PA      (SCRI ) SCRIPPS RES INST.
XX
PI      Barbas CF, Kadan M, Beerli R;
XX
DR      WPI; 2001-308618/32.
XX
PT      New fusion protein containing nucleotide-binding and ligand-binding
PT      domains, useful e.g. in gene therapy of cancer, provides ligand-activated
PT      control of gene expression.
XX
FS      Claim 25; Page 168-170; 218pp; English.
XX
CC      The invention relates to fusion protein comprising a nucleotide-binding
CC      domain (NBD), a ligand-binding domain (LBD) of an intracellular receptor
CC      (ICR) and a transcription regulating domain (TRD). NBD is a polydactyl
CC      zinc finger protein (ZFP), or a modular part of it, that interacts
```


CC specifically with a contiguous sequence of at least 3 nucleotides. The
CC fusion protein functions as a ligand-activated transcriptional regulator.
CC The fusion protein and the nucleic acid encoding it, are used to regulate
CC gene expression, particularly in gene therapy for treating malignant cell
CC proliferative diseases (e.g. colon cancer, prostate cancer, renal-cell
CC carcinoma) and non-malignant cell proliferative diseases (e.g. psoriasis,
CC pemphigus vulgaris, Behcet's syndrome and lipid histiocytosis). The
CC fusion protein and its DNA are also useful for treating diseases caused
CC by viruses in humans/plants, genetic and/or acquired diseases. The fusion
CC protein can be designed to target any selected gene (endogenous or
CC exogenous), and can be made to have different selectivity or specificity
CC for endogenous or exogenous ligands. The present sequence is 2C7LBDAS
CC construct encoding fusion protein comprising 2C7 zinc finger protein
CC (ZFP) which is a variant of murine Cys2-His2 ZFP Zif268, human oestrogen
CC receptor LBD fragment and VP16 transactivation domain (TA) from Herpes
CC simplex virus. The ZFP serves as NBD and VP16 TA domain functions as
CC transcription activator. The 2C7LBDAS construct is based on plasmid
CC pCDNA3.1 that comprises sequences from cytomegalovirus, bacteriophage T7
CC and simian virus 40 (SV40)

XX
SQ Sequence 6828 BP; 1595 A; 1816 C; 1746 G; 1681 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 4; Length 6828;
Best Local Similarity 100.0%; Pred. No. 3.3e-24;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAGCGCGCGCTCGAGTCTAGAGGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCT 60
Db 2432 AAGCGCGCGCTCGAGTCTAGAGGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCT 2491
QY 61 TCTAGTTGCCAGCCATCTGTGTTTGGCCCTCCCGGTC 101
Db 2492 TCTAGTTGCCAGCCATCTGTGTTTGGCCCTCCCGGTC 2532

Search completed: July 14, 2005, 07:01:43
Job time : 143.448 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 05:15:57 ; Search time 961.667 Seconds
(without alignments)
3997.736 Million cell updates/sec

Title: US-09-482-682-47_COPY_3944_4044

Perfect score: 101

Sequence: 1 aagcgccgctcgagtagtag.....gttgcccccctcccggtgcc 101

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsl1:*
9: gb_gsl2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	99	98.0	378	7	CF315931 HD--05-A1
2	91	90.1	605	7	CK719567 19817 Swo
3	84	83.2	400	7	CR850860 10869 Sto
4	68.2	67.5	132	9	CR074510 Forward s
5	68.2	67.5	286	9	CR083191 Forward s
6	66	65.3	295	7	CN78129 pgn2c.pk0
7	66	65.3	519	5	BM888450 Tml08 Hum
8	66	65.3	521	5	BM887817 TMS53 Hum
9	65.6	65.0	233	9	CR154962 Forward s
10	65	64.4	166	9	CR092687 Forward s
11	64.4	63.8	534	5	BM887701 TM304 Hum
12	64.2	63.6	329	9	CG632479 OST350781
13	63.4	62.8	600	5	BM887768 TM397 Hum
14	61.8	61.2	75	9	CR037248 Forward s
15	61.8	61.2	87	9	CR106833 Forward s
16	61.8	61.2	130	9	BM988352 Forward s
17	61	60.4	234	9	CR070494 Forward s
18	60.2	59.6	330	9	CR006502 Forward s
19	60	59.4	605	5	BM888562 TMM237 Hu
20	59.8	59.2	108	9	CR173214 Forward s
21	59.4	58.8	304	9	BM997931 Forward s
22	59	58.4	104	9	CR104210 Forward s
23	59	58.4	110	9	BM982981 Forward s
24	58.6	58.0	323	9	CR100521 Forward s

c 25	58.2	57.6	158	9	CR018574 Forward s
c 26	58.2	57.6	197	9	CR014355 Forward s
c 27	58	57.4	471	4	BM819796 K-EST0087
c 28	57.8	57.2	664	9	CR160587 Forward s
c 29	57.2	56.6	284	9	BM984480 Forward s
c 30	57	56.4	158	9	CR117924 Forward s
c 31	56.8	56.2	199	9	CR047320 Forward s
c 32	56.6	56.0	141	9	CR126132 Forward s
c 33	56.6	56.0	159	9	CR133954 Forward s
c 34	56.6	56.0	160	9	CR012517 Forward s
c 35	56.4	55.8	109	9	CR108493 Forward s
c 36	56	55.4	99	9	CR009197 Forward s
c 37	56	55.4	347	9	CR045655 Forward s
c 38	55.8	55.2	77	9	CR171087 Forward s
c 39	55.8	55.2	89	9	CR081749 Forward s
c 40	55.8	55.2	96	9	CR180417 Forward s
c 41	55.4	54.9	104	9	CR163542 Forward s
c 42	55.4	54.9	113	9	CR100912 Forward s
c 43	55.4	54.9	132	9	CR081810 Forward s
c 44	55.2	54.7	277	9	CR144680 Forward s
c 45	55	54.5	89	9	CR160035 Forward s

ALIGNMENTS

RESULT 1
CF315931
LOCUS 378 bp mRNA linear EST 15-AUG-2003
DEFINITION HD--05-A13.bl OSHDAC1-overexpressing transgenic rice plasmid cDNA library (HD) Oryza sativa (japonica cultivar-group) cDNA clone
HD--05-A13, mRNA sequence.
ACCESSION CF315931
VERSION CF315931.1 GI:33687692
KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 378)
AUTHORS Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C., Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
TITLE Large-scale Sequencing Analysis of Rice ESTs
JOURNAL Unpublished (2003)
COMMENT Contact: Nahm B.H.
Genomics and Genetics Institute, GreenGene Biotech Inc.; Division of Bioscience and Bioinformatics, Myongji University
Yongin, Kyeonggi, Korea
Tel: 82 31 330 6193
Fax: 82 31 321 6355
Email: bnhahm@gbio.com, bnhahm@bio.myongji.ac.kr.

FEATURES

source
1..378
Location/Qualifiers
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nackdong"
/db_xref="taxon:39947"
/clone="HD--05-A13"
/tissue_type="callus"
/dev_stage="proliferated callus on 2N6 media for 2 weeks"
/lab_host="E.coli DH10B"
/clone_lib="OSHDAC1-overexpressing transgenic rice plasmid cDNA library (HD)"
/notes="vector: PCR4-TOPO; Site 1: EcoRI; Callus was treated with ABA(20um) for 1hr. Oligo-capped mRNA was reverse transcribed and then used for PCR. mRNA was derived from rice Histone Deacetylase overexpression line."

ORIGIN

Query Match 98.0%; Score 99; DB 7; Length 378;
Best Local Similarity 100.0%; Pred. No. 2.5e-21;

```

Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GCGGCGCTCGAGTCTAGAGGCGCCGTTTAAACCGCTGATCAGCCTCGACTGTCCTTC 62
    |||
Db 79 GCGGCGCTCGAGTCTAGAGGCGCCGTTTAAACCGCTGATCAGCCTCGACTGTCCTTC 138
    |||

Qy 63 TAGTTGCCAGCCATCTGTTGTTTCCCTCCCTCCCGTGCC 101
    |||
Db 139 TAGTTGCCAGCCATCTGTTGTTTCCCTCCCTCCCGTGCC 177
    |||

RESULT 2
CK719567
LOCUS 19817 Swollen Stolon Solanum tuberosum cDNA, mRNA sequence. EST 10-FEB-2004
DEFINITION CK719567
ACCESSION CK719567
VERSION CK719567.1 GI:42511281
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 605)
AUTHORS Flinn,B., Rothwell,C., Sardana,R., Griffiths,R., Laque,M.,
DeKoeyer,D., Audy,P., Goyer,C., Li,X.-Q., Wang-Pruski,G. and
Regan,S.
TITLE Generation of ESTs from swollen stolon tissues of potato
JOURNAL Unpublished (2004)
COMMENT The Canadian Potato Genome Project - BioAtlantech
921 College Hill Rd, Fredericton, ON, E3B 6Z9, CANADA
Email: bflinn@bioatlantech.nb.ca
Clones can be requested from BioAtlantech via
bflinn@bioatlantech.nb.ca
Seq primer: T3.
FEATURES
    source
        Location/Qualifiers
            1..605
                /organism="Solanum tuberosum"
                /mol_type="mRNA"
                /cultivar="Shepody"
                /db_xref="taxon:4113"
                /tissue_type="Stolon"
                /lab_host="XL10-Gold"
                /clone_lib="Swollen Stolon"
                /note="Vector: pBluescript II SK(+); Site:1: EcoRI;
                Site 2: XhoI; supplier: Developmental series. Plants from
                pathogen-free Solanum tuberosum var. Shepody, clone 1756,
                nuclear stock were grown in a greenhouse under natural
                conditions. RNA was isolated from swollen stolon tissue,
                3-10mm in length, which was cut from the tip, to the base
                of swelling."
ORIGIN
Query Match 90.1%; Score 91; DB 7; Length 605;
Best Local Similarity 94.9%; Pred. No. 1e-18; Indels 0; Gaps 0;
Matches 94; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 GCGGCGCTCGAGTCTAGAGGCGCCGTTTAAACCGCTGATCAGCCTCGACTGTCCTTC 62
    |||
Db 393 GCGGCGCTCGAGTCTAGAGGCGCCGTTTAAACCGCTGATCAGCCTCGACTGTCCTTC 452
    |||

Qy 63 TAGTTGCCAGCCATCTGTTGTTTCCCTCCCTCCCGTGCC 101
    |||
Db 453 TAGTTGCCAGCCATCTGTTGTTTCCCTCCCTCCCGTGAC 491
    |||

RESULT 3
CK850860
LOCUS 10869 Stolon Solanum tuberosum cDNA, mRNA sequence. EST 08-MAR-2004
DEFINITION CK850860
ACCESSION CK850860.1 GI:45239470
VERSION

```

```

KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 400)
AUTHORS Flinn,B., Rothwell,C., Sardana,R., Griffiths,R., Laque,M., De
Koeyer,D., Audy,P., Goyer,C., Li,X.-Q., Wang-Pruski,G. and Regan,S.
TITLE Generation of ESTs from stolon tissues of potato
JOURNAL Unpublished (2004)
COMMENT Contact: Barry Flinn
The Canadian Potato Genome Project - BioAtlantech
921 College Hill Rd, Fredericton, ON, E3B 6Z9, CANADA
Email: bflinn@bioatlantech.nb.ca
Seq primer: T3.
FEATURES
    source
        Location/Qualifiers
            1..400
                /organism="Solanum tuberosum"
                /mol_type="mRNA"
                /cultivar="Shepody"
                /db_xref="taxon:4113"
                /tissue_type="Stolon"
                /lab_host="XL10-Gold"
                /clone_lib="Stolon"
                /note="Vector: pBluescript II SK(+); Site:1: EcoRI;
                Site 2: XhoI; supplier: Developmental series. Plants from
                pathogen-free Solanum tuberosum var. Shepody, clone 1756,
                nuclear stock were grown in a greenhouse under natural
                conditions. RNA was isolated from stolon tissue."
ORIGIN
Query Match 83.2%; Score 84; DB 7; Length 400;
Best Local Similarity 99.0%; Pred. No. 1.8e-16;
Matches 95; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 3 GCGGCGCTCGAGTCTAGAGGCGCCGTTTAAACCGCTGATCAGCCTCGACTGTCCTTC 62
    |||
Db 224 GCGGCGCTCGAGTCTAGAGGCGCCGTTTAAACCGCTGATCAGCCTCGACTGTCCTTC 283
    |||

Qy 63 TAGTTGCCAGCCA-TCTGTTGTTTGGCCCTCCCTCCCGG 97
    |||
Db 284 TAGTTGCCAGCCA-TCTGTTGTTTGGCCCTCCCGG 319
    |||

RESULT 4
CR074510/c
LOCUS 132 bp DNA linear GSS 05-JUL-2004
DEFINITION Forward strand read from insert in 3'HPRT insertion targeting and
chromosome engineering clone MHP255d22, genomic survey sequence.
ACCESSION CR074510
VERSION CR074510.1 GI:49808100
KEYWORDS GSS; genome survey sequence; MICER.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 132)
AUTHORS Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,I.,
Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
Rogers,J. and Bradley,A.
TITLE Direct Submission
JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES
    source
        Location/Qualifiers
            1..132
                /organism="Mus musculus"
                /mol_type="genomic DNA"
                /db_xref="taxon:10090"
                /clone="MHP255d22"
                /clone_lib="MHP"
ORIGIN

```

```

Query Match          67.5%; Score 68.2; DB 9; Length 132;
Best Local Similarity 90.1%; Pred. No. 1.9e-11;
Matches 73; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 21 AGGCGCCGTTAAACCCGCTGATCAGCTCGACTGTGCTTCTAGTTCGCCAGCCATCTGT 80
    |||||
Db 93 ACGACCCCATGATCGCGCTGATCAGCTCGACTGTGCTTCTAGTTCGCCAGCCATCTGT 34

Qy 81 TGTGTGCCCTCCCGCCGTGCC 101
    |||||
Db 33 TGTGTGCCCTCCCGCCGTGCC 13

RESULT 5
CR083191/c
LOCUS
DEFINITION
  CR083191 286 bp DNA linear GSS 05-JUL-2004
  Forward strand read from insert in 3'HPRT insertion targeting and
  chromosome engineering clone MHP263n05, genomic survey sequence.
ACCESSION
  CR083191.1 GI:49816780
VERSION
  GSS; genome survey sequence; MICER.
KEYWORDS
  Mus musculus (house mouse)
ORGANISM
  Mus musculus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  1 (Bases 1 to 286)
  Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
  Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
  Rogers,J. and Bradley,A.
  Direct Submission
  Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
  CB10 1SA, UK. http://www.sanger.ac.uk/MICER
  Location/Qualifiers
    1..286
      /organism="Mus musculus"
      /mol_type="genomic DNA"
      /db_xref="taxon:10090"
      /clone="MHP263n05"
      /clone_lib="MHP"

ORIGIN
source
Query Match          67.5%; Score 68.2; DB 9; Length 286;
Best Local Similarity 90.1%; Pred. No. 2.2e-11;
Matches 73; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 21 AGGCGCCGTTAAACCCGCTGATCAGCTCGACTGTGCTTCTAGTTCGCCAGCCATCTGT 80
    |||||
Db 88 ACGACCCCATGATCGCGCTGATCAGCTCGACTGTGCTTCTAGTTCGCCAGCCATCTGT 29

Qy 81 TGTGTGCCCTCCCGCCGTGCC 101
    |||||
Db 28 TGTGTGCCCTCCCGCCGTGCC 8

RESULT 6
CN778129
LOCUS
DEFINITION
  pgn2c.pk001.h10.f Chicken lymphoid cDNA library (pgn2c) Gallus
  gallus cDNA clone pgn2c.pk001.h10.f 3'end of pat.pk0008.d12 5',
  mRNA sequence.
ACCESSION
  CN778129
KEYWORDS
  EST.
SOURCE
  Gallus gallus (chicken)
  ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
  Phasianinae; Gallus.
  1 (Bases 1 to 295)
  Morgan,R.W. and Burnside,J.
  Chicken ESTs from lymphoid tissue- 3' sequence
  Unpublished (2004)
  Contact: Robin W. Morgan

Qy 81 TGTGTGCCCTCCCGCCGTGCC 101
    |||||
Db 28 TGTGTGCCCTCCCGCCGTGCC 8

RESULT 7
BM888450
LOCUS
DEFINITION
  TM108 Human Trabecular Meshwork cDNA library Homo sapiens cDNA
  clone 104447 5', mRNA sequence.
ACCESSION
  BM888450
VERSION
  BM888450.1 GI:19272194
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
  ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (Bases 1 to 519)
  Wirtz,M.K., Samples,J.R., Xu,H., Severson,T. and Acott,T.S.
  Expression Profile and Genome Location of cDNA Clones from an
  Infant Human Trabecular Meshwork Library
  Unpublished (2002)
  Contact: Wirtz MK
  Glaucoma Genetics Lab
  Oregon Health Sciences University
  3375 S.W. Terwilliger Blvd., Portland, OR 97201-4197, USA
  Tel: 503-494-4698
  Fax: 503-494-6875
  Email: wirtzm@ohsu.edu
  Seq primer: T7 Reverse
  Location/Qualifiers
    1..519
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="104447"
      /tissue_type="eye"
      /cell_type="trabecular meshwork"
      /dev_stage="2 week to 2 year old infants"
      /lab_host="TOP10P"
      /clone_lib="Human Trabecular Meshwork cDNA library"
      /note="Vector: pcDNA3; Site_1: EcoRI; Site_2: EcoRI; Human
      cDNA library made from mRNA isolated from trabecular
      meshwork cells established from eyes from 6 individuals,
      ages 2 weeks to 2 years. Cells were harvested at passages
      3 through 6. Invitrogen made a unidirectional cDNA library
      from the mRNA from the frozen cells using a pcDNA3 vector

```

```

University of Delaware
Townsend Hall, Newark, DE 19717, USA
Tel: 302-831-1341
Fax: 302 831-2822
Email: morgan@udel.edu, www.chickest.udel.edu.
Location/Qualifiers
  1..295
    /organism="Gallus gallus"
    /mol_type="mRNA"
    /db_xref="taxon:9031"
    /clone="pgn2c.pk001.h10.f 3'end of pat.pk0008.d12"
    /sex="Male and Female"
    /tissue_type="thymus, bursa, spleen, PBL, bone marrow"
    /lab_host="E.coli EMDH10B"
    /clone_lib="Chicken Lymphoid cDNA library (pgn2c)"
    /note="Vector: pCMVSPORT 6"

FEATURES
source
Query Match          65.3%; Score 66; DB 7; Length 295;
Best Local Similarity 93.2%; Pred. No. 1.1e-10;
Matches 69; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 28 GTTTAAACCCGCTGATCAGCTCGACTGTGCTTCTAGTTCGCCAGCCATCTGTGTTGC 87
    |||||
Db 1 GCTAGAGCTCGCTGATCAGCTCGACTGTGCTTCTAGTTCGCCAGCCATCTGTGTTGC 60

Qy 88 CCTCTCCCGCTGCC 101
    |||||
Db 61 CCTCTCCCGCTGCC 74

RESULT 7
BM888450
LOCUS
DEFINITION
  TM108 Human Trabecular Meshwork cDNA library Homo sapiens cDNA
  clone 104447 5', mRNA sequence.
ACCESSION
  BM888450
VERSION
  BM888450.1 GI:19272194
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
  ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (Bases 1 to 519)
  Wirtz,M.K., Samples,J.R., Xu,H., Severson,T. and Acott,T.S.
  Expression Profile and Genome Location of cDNA Clones from an
  Infant Human Trabecular Meshwork Library
  Unpublished (2002)
  Contact: Wirtz MK
  Glaucoma Genetics Lab
  Oregon Health Sciences University
  3375 S.W. Terwilliger Blvd., Portland, OR 97201-4197, USA
  Tel: 503-494-4698
  Fax: 503-494-6875
  Email: wirtzm@ohsu.edu
  Seq primer: T7 Reverse
  Location/Qualifiers
    1..519
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="104447"
      /tissue_type="eye"
      /cell_type="trabecular meshwork"
      /dev_stage="2 week to 2 year old infants"
      /lab_host="TOP10P"
      /clone_lib="Human Trabecular Meshwork cDNA library"
      /note="Vector: pcDNA3; Site_1: EcoRI; Site_2: EcoRI; Human
      cDNA library made from mRNA isolated from trabecular
      meshwork cells established from eyes from 6 individuals,
      ages 2 weeks to 2 years. Cells were harvested at passages
      3 through 6. Invitrogen made a unidirectional cDNA library
      from the mRNA from the frozen cells using a pcDNA3 vector

```

Qy	28	GTTTAAACCCGCTGATCAGCCTCGACTGTGCGCTTCTAGTTGCCAGCCATCTGTGTTTGC	87
Db	348	GCTAGAGCTCGCTGATCAGCCTCGACTGTGCTTCTAGTTGCCAGCCATCTGTGTTTGC	407
Qy	88	CCCTCCCCCGTGCC	101
Db	408	CCCTCCCCCGTGCC	421

Matches 71; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 21 AGGGCCCTTTAAACCCGCTGATCAGCTCGACTGTGCTTCTAGTTCGCCAGCCATCTGT 80
 Db 156 ACGACCCCATGATCGCGCTGATCAGCTCGAGTGTGCTTCTAGTTCGCCAGCCATCTGT 97

Qy 81 TGTGTGCCCTCCCGTGCC 101
 Db 96 TGTGTGCCCTCCCGTGCC 76

RESULT 11
 BM887701 534 bp mRNA linear EST 08-MAR-2002
 LOCUS TM304 Human Trabecular Meshwork cDNA library Homo sapiens cDNA
 DEFINITION clone 107917 5', mRNA sequence.

ACCESSION BM887701
 VERSION BM887701.1 GI:19271430
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 534)
 Wirtz,M.K., Samples,J.R., Xu,H., Severson,T. and Acott,T.S.
 Expression Profile and Genome Location of cDNA Clones from an
 Infant Human Trabecular Meshwork Library
 Unpublished (2002)
 JOURNAL Contact: Wirtz MK
 COMMENT Glaucoma Genetics Lab
 Oregon Health Sciences University
 3375 S.W. Terwilliger Blvd., Portland, OR 97201-4197, USA
 Tel: 503-494-4698
 Fax: 503-494-6875
 Email: wirtzm@ohsu.edu
 Seq primer: 17 Reverse.
 Location/Qualifiers
 1..534
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="107917"
 /tissue_type="eye"
 /cell_type="trabecular meshwork"
 /dev_stage="2 week to 2 year old infants"
 /lab_host="TOP10P"
 /clone_lib="Human Trabecular Meshwork cDNA library"
 /note="Vector: pcDNA3; Site 1: EcoRI; Site 2: EcoRI; Human
 cDNA library made from mRNA isolated from trabecular
 meshwork cells established from eyes from 6 individuals,
 ages 2 weeks to 2 years. Cells were harvested at passages
 3 through 6. Invitrogen made a unidirectional cDNA library
 from the mRNA from the frozen cells using a pcDNA3 vector
 and TPO10P," host cells."

ORIGIN

Query Match 63.8%; Score 64.4; DB 5; Length 534;
 Best Local Similarity 91.9%; Pred. No. 4.2e-10;
 Matches 68; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 28 GTTTAAACCGCTGATCAGCTCGACTGTGCTTCTAGTTCGCCAGCCATCTGTGTTGC 87
 Db 445 GCTAGAGTTCGCTGATCAGCTCGACTGTGCTTCTAGTTCGCCAGCCATCTGTGTTGC 504

Qy 88 CCTCCCTCCCGTGCC 101
 Db 505 CCTCCCTCCCGTGCC 518

RESULT 12
 CG632479 329 bp mRNA linear GSS 02-OCT-2003
 LOCUS OST350781 Mus musculus 129Sv/Ev Mus musculus cDNA clone OST350781,
 DEFINITION

mRNA sequence.
 CG632479 GI:37456328
 CG632479.1
 GSS.
 Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 329)
 zambrowicz,B.P., Abuin,A., Ramirez-Solis,R., Richter,L.J.,
 Piggott,J., BeltrandelRio,H., Buxton,E.C., Edwards,J., Finch,R.A.,
 Fiddie,C.J., Gupta,A., Hansen,G., Hu,Y., Huang,W., Jaing,C.,
 Key,B.W. Jr., Kipp,P., Kohlhauff,B., Ma,Z.-Q., Markesich,D.,
 Payne,R., Potter,D.G., Qian,N., Shaw,J., Schrick,J., Shi,Z.-Z.,
 Sparks,M.J., Van Sligtenhorst,I., Vogel,P., Walke,W., Xu,N.,
 Zhu,Q., Person,C. and Sands,A.T.
 Wnk1 kinase deficiency lowers blood pressure in mice: a gene-trap
 screen to identify potential targets for therapeutic intervention
 Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)
 Contact: Zambrowicz BP
 OmniBank
 Lexicon Genetics Incorporated
 4000 Research Forest Drive, The Woodlands, TX 77381, USA
 Email: materials@lexgen.com
 Gene trap sequence tag generated by 3' RACE from mouse ES cells as
 described in Zambrowicz et al (Nature. 1998 Apr 9;392(6676):608-11)
 Class: Gene Trap.
 Location/Qualifiers
 1..329
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="129Sv/Ev"
 /db_xref="taxon:10090"
 /clone="OST350781"
 /cell_type="embryonic stem cell"
 /clone_lib="Mus musculus 129Sv/Ev"

FEATURES
 source

ORIGIN

Query Match 63.6%; Score 64.2; DB 9; Length 329;
 Best Local Similarity 95.7%; Pred. No. 4.5e-10;
 Matches 66; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 33 AACCCGCTGATCAGCTCGACTGTGCTTCTAGTTCGCCAGCCATCTGTGTTGCCCTC 92
 Db 185 AGCTCGCTGATCAGCTCGACTGTGCTTCTAGTTCGCCAGCCATCTGTGTTGCCCTC 244

Qy 93 CCCCCTGCC 101
 Db 245 CCCCCTACC 253

RESULT 13
 BM887768 600 bp mRNA linear EST 08-MAR-2002
 LOCUS TM397 Human Trabecular Meshwork cDNA library Homo sapiens cDNA
 DEFINITION clone 119752 5', mRNA sequence.

ACCESSION BM887768
 VERSION BM887768.1 GI:19271512
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 600)
 Wirtz,M.K., Samples,J.R., Xu,H., Severson,T. and Acott,T.S.
 Expression Profile and Genome Location of cDNA Clones from an
 Infant Human Trabecular Meshwork Library
 Unpublished (2002)
 JOURNAL Contact: Wirtz MK
 COMMENT Glaucoma Genetics Lab
 Oregon Health Sciences University
 3375 S.W. Terwilliger Blvd., Portland, OR 97201-4197, USA
 Tel: 503-494-4698

Fax: 503-494-6875
Email: wirtzm@ohsu.edu
Seq primer: T7 Reverse
High quality sequence stop: 400.

FEATURES

source
1. .600
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="119752"
/tissue_type="eye"
/cell_type="trabecular meshwork"
/dev_stage="2 week to 2 year old infants"
/lab_host="TOP10P"
/clone_lib="Human Trabecular Meshwork cDNA library"
/note="vector: pcDNA3; Site 1: EcoRI; Site 2: EcoRI; Human cDNA library made from mRNA isolated from trabecular meshwork cells established from eyes from 6 individuals, ages 2 weeks to 2 years. Cells were harvested at passages 3 through 6. Invitrogen made a unidirectional cDNA library from the mRNA from the frozen cells using a pcDNA3 vector and TPO10P, host cells."

ORIGIN

Query Match 62.8%; Score 63.4; DB 5; Length 600;
Best Local Similarity 90.5%; Pred. No. 9e-10; Indels 0; Gaps 0;
Matches 67; Conservative 0; Mismatches 7;

Qy 28 GTTTAAACCCGCTGATCAGCTCGACTGTGCTTCTAGTTGCCAGCCATCTGTTGTTGC 87
Db 516 GCTAGAGTCGCTGATCAGCTCGACTGTGCTTCTAGTTGCCAGCCATCTGTTGTTGC 575

Qy 88 CCCTCCCGGTGCC 101
Db 576 CCCTCCCGGTGCC 589

RESULT 14

CR037248/c
LOCUS
DEFINITION Forward strand read from insert in 3'HPRT insertion targeting and chromosome engineering clone MHP31109, genomic survey sequence.

ACCESSION CR037248
VERSION CR037248.1 GI:49770303
KEYWORDS GSS; genome survey sequence; MICER.
SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L., Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y., Rogers,J. and Bradley,A.

TITLE

Direct Submission

JOURNAL

Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,

FEATURES

CB10 1SA, UK. http://www.sanger.ac.uk/MICER
Location/Qualifiers
1. .75
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="MHP31109"
/clone_lib="MHP"

ORIGIN

Query Match 61.2%; Score 61.8; DB 9; Length 75;
Best Local Similarity 96.9%; Pred. No. 2.1e-09;
Matches 63; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 37 CGCTGATCAGCTCGACTGTGCTTCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCCC 96
Db 74 CGCTGATCAGCTCGACTGTGCTTCTAGTTGCCAGCCATCTGTTGTTGCCCTCCCCC 15

Qy 97 GTGCC 101
Db 14 GTGCC 10

RESULT 15

CR106833/c
LOCUS
DEFINITION Forward strand read from insert in 3'HPRT insertion targeting and chromosome engineering clone MHP37h18, genomic survey sequence.

ACCESSION CR106833
VERSION CR106833.1 GI:49854244
KEYWORDS GSS; genome survey sequence; MICER.
SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L., Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y., Rogers,J. and Bradley,A.

TITLE

Direct Submission

JOURNAL

Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,

CB10 1SA, UK. http://www.sanger.ac.uk/MICER

FEATURES

source
1. .87
Location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="MHP37h18"
/clone_lib="MHP"

ORIGIN

Query Match 61.2%; Score 61.8; DB 9; Length 87;
Best Local Similarity 85.2%; Pred. No. 2.1e-09;
Matches 69; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 21 AGGGCCGTTAAACCCGCTGATCAGCTCGACTGTGCTTCTAGTTGCCAGCCATCTGT 80
Db 86 ACAGCCCATGATCGATCGCATGATCAGCCTCAACTCTGCTTCTAGTTGCCAGCCATCTGT 27

Qy 81 TGTTTGGCCCTCCCGGTGCC 101
Db 26 TGTTTGGCCCTCCCGGTGCC 6

Search completed: July 14, 2005, 23:23:05
Job time : 961.667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:39:07 ; Search time 756.618 Seconds
(without alignments)
6468.225 Million cell updates/sec

Title: US-09-482-682-47_COPY_7889_7989

Perfect score: 101

Sequence: 1 aggttattgtctcatgacg.....gaaaagtgcacctgacgtc 101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*
2: gb_hlg.*
3: gb_in.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_bts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	101	100.0	142	6	AR356490 Sequence
C 2	101	100.0	142	6	AR538046 Sequence
C 3	101	100.0	228	6	E00019 DNA coding
C 4	101	100.0	240	1	M10199 Plasmid pMM
C 5	101	100.0	251	6	E00018 DNA coding
C 6	101	100.0	251	6	I01644 Sequence 1
C 7	101	100.0	344	11	LI8624 Human chrom
C 8	101	100.0	400	6	BD195256 Nucleotid
C 9	101	100.0	456	6	E00892 Synthetic D
C 10	101	100.0	456	6	E01156 DNA fragmen
C 11	101	100.0	456	6	E01274 DNA encodin
C 12	101	100.0	456	6	E01302 DNA encodin
C 13	101	100.0	466	6	AX260098 Sequence
C 14	101	100.0	573	6	AX260150 Sequence
C 15	101	100.0	693	6	A43586 Sequence 11
C 16	101	100.0	693	6	AR116755 Sequence
C 17	101	100.0	998	1	AY559171 Pseudomon
C 18	101	100.0	1011	1	SMTEMAQGE
C 19	101	100.0	1012	2	CEC11F10

20	101	100.0	1014	4	CFAJ4121	AJ224121	Canis fam
C 21	101	100.0	1027	1	AY589493	AY589493	Escherich
C 22	101	100.0	1040	1	AY538698	AY538698	Serratia
C 23	101	100.0	1040	1	AY538700	AY538700	Serratia
C 24	101	100.0	1040	1	AY538701	AY538701	Serratia
C 25	101	100.0	1040	1	AY538702	AY538702	Serratia
C 26	101	100.0	1041	1	AY538699	AY538699	Serratia
C 27	101	100.0	1042	1	AY394610	AY394610	Klebsiell
C 28	101	100.0	1042	1	ECO308558	AJ08558	Escherich
C 29	101	100.0	1044	1	AY392531	AY392531	Streptoco
C 30	101	100.0	1044	1	AY452662	AY452662	Streptoco
C 31	101	100.0	1054	1	AF104441	AF104441	Klebsiell
C 32	101	100.0	1054	1	AF104442	AF104442	Escherich
C 33	101	100.0	1058	6	I03356	I03356	Sequence 4
C 34	101	100.0	1064	1	AY628199	AY628199	Escherich
C 35	101	100.0	1069	1	AF535127	AF535127	Klebsiell
C 36	101	100.0	1069	1	AY243512	AY243512	Klebsiell
C 37	101	100.0	1071	1	AY628175	AY628175	Escherich
C 38	101	100.0	1072	1	AY101764	AY101764	Klebsiell
C 39	101	100.0	1073	6	AR371489	AR371489	Sequence
C 40	101	100.0	1073	6	AX195443	AX195443	Sequence
C 41	101	100.0	1075	1	AY729027	AY729027	Proteus m
C 42	101	100.0	1075	1	PATN1PN2	X54606	Pseudomonas
C 43	101	100.0	1075	1	PATN2PN1B	X54607	Pseudomonas
C 44	101	100.0	1075	1	PATN3PN1A	X54604	Pseudomonas
C 45	101	100.0	1080	1	AF027199	AF027199	Klebsiell

ALIGNMENTS

RESULT 1	AR356490/c	AR356490	142 bp	DNA	linear	PAT 17-AUG-2003
LOCUS	Sequence 2608 from patent US 6593114.					
DEFINITION	Sequence 2608 from patent US 6593114.					
ACCESSION	AR356490					
VERSION	AR356490.1	GI:33762574				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	1 (bases 1 to 142)					
AUTHORS	Kunsch,C.A., Choi,G.H., Barash,S., Dillon,P.J., Fannon,M.R. and Rosen,C.A.					
TITLE	Staphylococcus aureus polynucleotides and sequences					
JOURNAL	Patent: US 6593114-A 2608 15-JUL-2003;					
FEATURES	Location/Qualifiers					
source	1..142					
	/organism="unknown"					
	/mol_type="genomic DNA"					
ORIGIN						

Query Match 100.0%; Score 101; DB 6; Length 142;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGCGTTATTTCTCATGCGGATACATATTGTGAATGTTTACAAAAATAACAATAG 60
|||||
Db 107 AGCGTTATTTCTCATGCGGATACATATTGTGAATGTTTACAAAAATAACAATAG 48
|||||

Qy 61 GGGTTCGCGCACATTTTCCCGAAAAGTGCACCTGACGTC 101
|||||
Db 47 GGGTTCGCGCACATTTTCCCGAAAAGTGCACCTGACGTC 7
|||||

RESULT 2	AR538046/c	AR538046	142 bp	DNA	linear	PAT 08-OCT-2004
LOCUS	Sequence 2608 from patent US 6737248.					
DEFINITION	Sequence 2608 from patent US 6737248.					
ACCESSION	AR538046					
VERSION	AR538046.1	GI:53929263				
KEYWORDS	Unknown.					
SOURCE	Unknown.					

```
ORGANISM      Unknown.
REFERENCE      Unclassified.
AUTHORS        1 (bases 1 to 142)
                Kunsch,C.A., Choi,G.A., Barash,S.C., Dillon,P.J., Fannon,M.R. and
                Rosen,C.A.
TITLE          Staphylococcus aureus polynucleotides and sequences
JOURNAL        Patent: US 6737248-A 2608 18-MAY-2004;
FEATURES       Location/Qualifiers
                source
                1..142
                /organism="unknown"
                /mol_type="genomic DNA"

ORIGIN
Query Match      100.0%; Score 101; DB 6; Length 142;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTATTAGAAAAATAAACAAATAG 60
Db 107 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTATTAGAAAAATAAACAAATAG 48

Qy 61 GGGTTCGGCGACATTTCCCGGAAAAGTGCCACCTGACGTC 101
Db 47 GGGTTCGGCGACATTTCCCGGAAAAGTGCCACCTGACGTC 7

RESULT 3
E00019/c
LOCUS          DNA coding for Escherichia coli penicillinase.
DEFINITION     E00019
ACCESSION      E00019
VERSION        E00019.1 GI:2168327
KEYWORDS       JP 1981154999-A/2.
SOURCE         Escherichia coli
ORGANISM       Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE      1 (bases 1 to 228)
AUTHORS        Uotutaa,G. and Karen,T.
TITLE          SYNTHESIS OF SELECTIVE AGED PROTEIN OR POLYPEPTIDE IN BACTERIA
JOURNAL        Patent: JP 1981154999-A 2 30-NOV-1981;
COMMENT        UNIV HARVARD
               OS Escherichia coli
               PN JP 1981154999-A/2
               PD 30-NOV-1981
               PF 09-APR-1981 JP 1981052488
               PR 11-APR-1980 US 80 139225
               PI UORUTAA GIRUBATO, KAREN TARUMATSUJI
               PC C12P21/00,C07H21/00,C12N1/00,C12N15/00//C12R1/19; CC
               strandedness: Double;
               CC topology: Linear;
               CC anti-sense: No;
               CC *source: clone=pKT218;
               FH Key Location/Qualifiers
               FH CDS 210..>228
               FT /product='E.coli penicillinase'.

FEATURES       source
               1..228
               Location/Qualifiers
               /organism="Escherichia coli"
               /mol_type="genomic DNA"
               /db_xref="taxon:562"

ORIGIN
Query Match      100.0%; Score 101; DB 6; Length 228;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTATTAGAAAAATAAACAAATAG 60
Db 175 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTATTAGAAAAATAAACAAATAG 116

Qy 61 GGGTTCGGCGACATTTCCCGGAAAAGTGCCACCTGACGTC 101
```

```
Db 115 GGGTTCGGCGACATTTCCCGGAAAAGTGCCACCTGACGTC 75

RESULT 4
PMOENDO/c
LOCUS          240 bp DNA linear BCT 26-APR-1993
DEFINITION     Plasmid pMM110 region of endo VII cleavage sites near cruciform
                structures.
ACCESSION      M10199
VERSION        M10199.1 GI:150826
KEYWORDS       Plasmid pMM110
SOURCE         Plasmid pMM110
ORGANISM       Plasmid pMM110
other sequences; plasmids.
REFERENCE      1 (bases 1 to 240)
AUTHORS        Kemper,B., Jensch,F., von Depka-Prondzynski,M., Fritz,H.J.,
                Borgmeyer,U. and Mizuuchi,K.
TITLE          Resolution of Holliday structures by endonuclease VII as observed
                in interactions with cruciform DNA
JOURNAL        Cold Spring Harb. Symp. Quant. Biol. 49, 815-825 (1984)
MEDLINE        85153063
PUBMED         6397324
COMMENT        Original source text: Plasmid pMM110 DNA.
FEATURES       source
               1..240
               Location/Qualifiers
               /organism="Plasmid pMM110"
               /mol_type="genomic DNA"
               /db_xref="taxon:2599"
               /plasmid="Plasmid pMM110"

ORIGIN          Unreported.
Query Match      100.0%; Score 101; DB 1; Length 240;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTATTAGAAAAATAAACAAATAG 60
Db 151 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTATTAGAAAAATAAACAAATAG 92

Qy 61 GGGTTCGGCGACATTTCCCGGAAAAGTGCCACCTGACGTC 101
Db 91 GGGTTCGGCGACATTTCCCGGAAAAGTGCCACCTGACGTC 51

RESULT 5
E00018/c
LOCUS          DNA coding for Escherichia coli penicillinase.
DEFINITION     E00018
ACCESSION      E00018
VERSION        E00018.1 GI:2168326
KEYWORDS       JP 1981154999-A/1.
SOURCE         Escherichia coli
ORGANISM       Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE      1 (bases 1 to 251)
AUTHORS        Uotutaa,G. and Karen,T.
TITLE          SYNTHESIS OF SELECTIVE AGED PROTEIN OR POLYPEPTIDE IN BACTERIA
JOURNAL        Patent: JP 1981154999-A 1 30-NOV-1981;
COMMENT        UNIV HARVARD
               OS Escherichia coli
               PN JP 1981154999-A/1
               PD 30-NOV-1981
               PF 09-APR-1981 JP 1981052488
               PR 11-APR-1980 US 80 139225
               PI UORUTAA GIRUBATO, KAREN TARUMATSUJI
               PC C12P21/00,C07H21/00,C12N1/00,C12N15/00//C12R1/19; CC
               strandedness: Double;
               CC topology: Linear;
               CC anti-sense: No;
               CC fragment_type: N-Terminal Fragment;
               CC *source: clone=pKT241;
```

FH Key Location/Qualifiers
FT CDS 210..>252
FT /product='E.coli penicillinase' FT
TATA_signal 190..196
Location/Qualifiers
1..251
/organism='Escherichia coli'
/mol_type='genomic DNA'
/db_xref='taxon:562'

FEATURES
source
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 251;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAG 60
|||||
DB 175 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAG 116
|||||

QY 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
|||||
DB 115 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 75
|||||

RESULT 6
I01644/c
LOCUS 251 bp ss-DNA linear PAT 18-MAY-1993
DEFINITION Sequence 1 from Patent US 4338397.
ACCESSION I01644
VERSION I01644.1 GI:267685
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 251)
AUTHORS Gilbert,W. and Talmadge,K.
TITLE Mature protein synthesis
JOURNAL Patent: US 4338397-A 1 06-JUL-1982;
President and Fellows of Harvard College; Cambridge, MA

FEATURES
source
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 251;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAG 60
|||||
DB 175 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAG 116
|||||

QY 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
|||||
DB 115 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 75
|||||

RESULT 7
HUMUT5345
LOCUS 344 bp DNA linear STS 26-JUL-1993
DEFINITION Human chromosome 8 STS UT5345, sequence tagged site.
ACCESSION L18624
VERSION L18624.1 GI:308338
KEYWORDS STS; PCR primer; STS sequence; microsatellite marker;
microsatellite repeat; repeat polymorphism; sequence tagged site.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 344)
AUTHORS Gerken,S.C., Matsunami,N., Lawrence,E., Carlson,M., Moore,M.,

Ballard,L., Melis,R., Robertson,M., Bradley,P., Elsner,T.,
Tingey,A., Rodriguez,P., Albertsen,H., Lalouel,J.-M. and White,R.
Genetic and physical mapping of simple sequence repeat containing
sequence tagged sites from the human genome
Unpublished (1993)
Original source text: Homo sapiens DNA.
Submitted by: Utah Center for Human Genome Research University of
Utah, Dept. of Human Genetics
2160 Eccles Institute of Human Genetics
Salt Lake City, UT 84112
e-mail: sts@corona.med.utah.edu
Primer A: GAGCAAAAACAGGAGGCAAAATGC
Primer B: TTCGGGAATGTCCCGGAAC
32P-label: B Primer
PCR Profile:
Initial Denaturation: 94C 300sec
PCR Cycles: 30
Denaturation: 94C 10sec
Annealing: 60C 10sec
Extension: 72C 20sec
Mg++: 2mM
Gel: Acrylamide 7%, Formamide 32%, Urea 34%
Alleles: 2.

FEATURES
Location/Qualifiers
1..344
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'
/map='8'
36..224
/standard_name='STS UT5345'
36..60
complement (202..224)

STS
primer_bind
primer_bind
ORIGIN
Query Match 100.0%; Score 101; DB 11; Length 344;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAG 60
|||||
DB 141 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAAATAG 200
|||||

QY 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
|||||
DB 201 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 241
|||||

RESULT 8
BD195256/c
LOCUS 400 bp DNA linear PAT 17-JUL-2003
DEFINITION Nucleotide sequence of Escherichia coli pathogenicity islands.
ACCESSION BD195256
VERSION BD195256.1 GI:33005021
KEYWORDS JP 2002513277-A/43.
SOURCE unidentified
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 400)
AUTHORS Dillion,P.J., Choi,G.H. and Welch,R.A.
TITLE Nucleotide sequence of Escherichia coli pathogenicity islands
JOURNAL Patent: JP 2002513277-A 43 08-MAY-2002;
HUMAN GENOME SCIENCES INC,WISCONSIN ALUMNI RESEARCH FOUNDATION

COMMENT
OS Unidentified
PN JP 2002513277-A/43
PD 08-MAY-2002
PF 21-NOV-1997 JP 1998523916
PR 22-NOV-1996 US 60/031626,14-OCT-1997 US 60/061953 PI
PATRICK J DILLON,GIL H CHOI,RODNEY A WELCH
PC C12N15/11,C12N15/63,C07K16/12,G01N33/569,G06F17/30,G11B7/00 CC
Strandedness: Double;
CC Topology: Linear;
CC Nucleotide sequence of Escherichia coli pathogenicity islands

```

FH Key Location/Qualifiers
FT source 1..400
FT /organism='Unidentified'.

FEATURES
    source
        1..400
        /organism='unidentified'
        /mol_type='genomic DNA'
        /db_xref='taxon:32644'

ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 400;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 60
    |||||||
Db 165 AGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 106
    |||||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGACGTC 101
    |||||||
Db 105 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGACGTC 65
    |||||||

RESULT 9
E00892/c
LOCUS E00892 456 bp DNA linear PAT 29-SEP-1997
DEFINITION Synthetic DNA encoding fused polypeptide between E coli
beta-lactamase and human beta-urogastrone.
ACCESSION E00892
VERSION E00892.1 GI:2169153
KEYWORDS JP 1986149089-A/1.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 456)
AUTHORS Okai,H., Momota,Y., Kumakura,T., Tochifusa,N., Kitazawa,T.,
Ojida,K. and Matsushiro,A.
TITLE POLYPEPTIDE SECRETION DEVELOPMENT VECTOR, MICROORGANISM TRANSFORMED
WITH SAME, AND PRODUCTION OF POLYPEPTIDE WITH MICROORGANISM
JOURNAL Patent: JP 1986149089-A 1 07-JUL-1986;
EARTH CHEM CORP LTD
COMMENT OS Artificial gene
OC Artificial sequence; Genes.
PN JP 1986149089-A/1
PD 07-JUL-1986
PF 21-DEC-1984 JP 1984271206
PI OKAI HIDEO, MOMOTA YUTAKA, KUMAKURA TAKESHI,
PI TOCHIFUSA NORIYUKI,
PI KITAZAWA TOSHIKI, OJIDA KAZUHIRO, MATSUSHIRO AIZO PC
C12N15/00,C12N1/20,C12P21/00,(C12N1/20,C12R1:19),(C12P21/00,PC
C12R1:19);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: strain=HB101;
CC *source: clone=pVG201;
CC Feature is identified by experimental;
FH Key Location/Qualifiers
FH promoter 125..170
FH of beta-lactamase
FT RBS 200..203
FT CDS 209..438
FT /product='beta-urogastrone precursor' FT
FT sig_peptide 209..277
FT /product='signal peptide of beta-lactonase' FT
FT mat_peptide 278..435
FT /product='beta-urogastrone mature peptide'.
FT Location/Qualifiers
FT 1..456
FT /organism='synthetic construct'
FT /mol_type='genomic DNA'

FEATURES
    source
        1..456
        /organism='synthetic construct'
        /mol_type='genomic DNA'

```

```

ORIGIN
/db_xref='taxon:32630'

Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 60
    |||||||
Db 173 AGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 114
    |||||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGACGTC 101
    |||||||
Db 113 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGACGTC 73
    |||||||

RESULT 10
E01156/c
LOCUS E01156 456 bp DNA linear PAT 29-SEP-1997
DEFINITION DNA fragment which secretes beta urogastrone.
ACCESSION E01156
VERSION E01156.1 GI:2169415
KEYWORDS JP 1987083890-A/1.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 456)
AUTHORS Yoshikawa,K., Momota,Y., Kajifusa,N., Koide,T. and Okai,H.
TITLE POLYPEPTIDE SECRETION EXPRESSION VECTOR, MICROORGANISM TRANSFORMED
BY SAID VECTOR AND PRODUCTION OF POLYPEPTIDE USING SAID
JOURNAL Patent: JP 1987083890-A 1 17-APR-1987;
EARTH CHEM CORP LTD
COMMENT OS Artificial gene
OC Artificial sequence; Genes.
PN JP 1987083890-A/1
PD 17-APR-1987
PF 09-OCT-1985 JP 1985225393
PI YOSHIKAWA KAZUTOSHI, MOMOTA YUTAKA, KAJIFUSA NORIYUKI, PI
KOIDE TAKAO,
PI OKAI HIDEO
PC C12N15/00,C12N1/20,C12P21/00,(C12N1/20,C12R1:19),(C12N1/20,PC
C12R1:125);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: clone=pUG201;
FH Key Location/Qualifiers
FH promoter 125..170
FH /note='beta lactamase promoter' FT RBS
FT CDS 209..439
FT /product='beta urogastrone'
FT sig_peptide 209..277
FT mat_peptide 278..436
FT /product='beta urogastrone'.
FH Key Location/Qualifiers
FH 1..456
FH /organism='synthetic construct'
FH /mol_type='genomic DNA'
FH /db_xref='taxon:32630'

ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 60
    |||||||
Db 173 AGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 114
    |||||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGACGTC 101

```

```
Db 113 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 73
|||||
RESULT 11
E01274/c
LOCUS
DEFINITION
E01274
ACCESSION
E01274.1 GI:2169533
KEYWORDS
JP 1987179398-A/1.
SOURCE
synthetic construct
ORGANISM
other sequences; artificial sequences.
REFERENCE
1 (bases 1 to 456)
AUTHORS
Okai,H., Kumakura,T., Kawamoto,S., Adachi,S., Matsubara,A.,
Ojida,K., Yano,M., Mihara,S., Matsushiro,A. and Yanaihara,N.
TITLE
PRODUCTION OF BETA-UROGASTRONE
JOURNAL
EARTH CHEM CORP LTD
COMMENT
OC Artificial gene
OS Homo sapiens
PN JP 1987179398-A/1
PD 06-AUG-1987
PF 31-JAN-1986 JP 1986021032
PI OKAI HIDEO, KUMAKURA TAKESHI, KAWAMOTO SHOICHI, ADACHI SHOJI,
MATSUBARA AKIMASA, OJIDA KAZUhide, YANO MAKI, MIHARA SHIGERU,
MATSUSHIRO AIZO, YANAIHARA NOBORU
PC C12P21/00,C12N15/00,(C12P21/00,C12R1:91);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No; Location/Qualifiers
FH Key
FH Promoter 125..170
FT RBS 200..203
FT sig_peptide 209..277
FT mat_peptide 278..436
FT /product='beta-urogastron'
FT CDS 209..439
FT /product='beta-urogastron'.
FEATURES
source
1..456
/organism='synthetic construct'
/mol_type='genomic DNA'
/db_xref='taxon:32630'
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20; Mismatches 0; Indels 0; Gaps 0;
Matches 101; Conservative 0;
QY 1 AGGGTTATTGTCATGCGGATACATATTTGAATGTTAGAAAAATAAACAATAG 60
|||||
Db 173 AGGGTTATTGTCATGCGGATACATATTTGAATGTTAGAAAAATAAACAATAG 114
|||||
QY 61 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 101
|||||
Db 113 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 73
|||||
RESULT 13
E01274/c
LOCUS
DEFINITION
E01274
ACCESSION
E01274.1 GI:16509129
KEYWORDS
Drosophila melanogaster (fruit fly)
SOURCE
Drosophila melanogaster
ORGANISM
Drosophila melanogaster
REFERENCE
1
AUTHORS
Deak,P., Glover,D.M. and Midgley,C.
TITLE
Cell cycle progression proteins
JOURNAL
Patent: WO 0172774-A 60 04-OCT-2001;
Cyclacel Limited (GB)
FEATURES
source
1..466
/organism='Drosophila melanogaster'
/mol_type='unassigned DNA'
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20; Mismatches 0; Indels 0; Gaps 0;
Matches 101; Conservative 0;
QY 1 AGGGTTATTGTCATGCGGATACATATTTGAATGTTAGAAAAATAAACAATAG 60
|||||
Db 173 AGGGTTATTGTCATGCGGATACATATTTGAATGTTAGAAAAATAAACAATAG 114
|||||
QY 61 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 101
|||||
Db 113 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 73
|||||
RESULT 13
E01302/c
LOCUS
DEFINITION
E01302
ACCESSION
E01302.1 GI:2169561
KEYWORDS
JP 1987190083-A/1.
SOURCE
synthetic construct
ORGANISM
other sequences; artificial sequences.
REFERENCE
1 (bases 1 to 456)
AUTHORS
Okai,H., Kumakura,T., Kawamoto,S., Adachi,S., Matsubara,A.,
Ojida,K., Yano,M., Mihara,S., Matsushiro,A. and Yanaihara,N.
TITLE
PRODUCTION OF BETA-UROGASTRONE
JOURNAL
EARTH CHEM CORP LTD
COMMENT
OC Artificial gene
OS Homo sapiens
PN JP 1987179398-A/1
PD 06-AUG-1987
PF 31-JAN-1986 JP 1986021032
PI OKAI HIDEO, KUMAKURA TAKESHI, KAWAMOTO SHOICHI, ADACHI SHOJI,
MATSUBARA AKIMASA, OJIDA KAZUhide, YANO MAKI, MIHARA SHIGERU,
MATSUSHIRO AIZO, YANAIHARA NOBORU
PC C12P21/00,C12N15/00,(C12P21/00,C12R1:91);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No; Location/Qualifiers
FH Key
FH Promoter 125..170
FT RBS 200..203
FT sig_peptide 209..277
FT mat_peptide 278..436
FT /product='beta-urogastron'
FT CDS 209..439
FT /product='beta-urogastron'.
FEATURES
source
1..456
/organism='synthetic construct'
/mol_type='genomic DNA'
/db_xref='taxon:32630'
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20; Mismatches 0; Indels 0; Gaps 0;
Matches 101; Conservative 0;
QY 1 AGGGTTATTGTCATGCGGATACATATTTGAATGTTAGAAAAATAAACAATAG 60
|||||
Db 173 AGGGTTATTGTCATGCGGATACATATTTGAATGTTAGAAAAATAAACAATAG 114
|||||
QY 61 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 101
|||||
Db 113 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 73
|||||
RESULT 12
E01302/c
LOCUS
DEFINITION
E01302
ACCESSION
E01302.1 GI:2169561
KEYWORDS
JP 1987190083-A/1.
SOURCE
synthetic construct
ORGANISM
other sequences; artificial sequences.
REFERENCE
1 (bases 1 to 456)
AUTHORS
Okai,H., Kumakura,T., Kawamoto,S., Adachi,S., Matsubara,A.,
Ojida,K., Yano,M., Mihara,S., Matsushiro,A. and Yanaihara,N.
TITLE
PRODUCTION OF BETA-UROGASTRONE
JOURNAL
EARTH CHEM CORP LTD
COMMENT
OC Artificial gene
OS Homo sapiens
PN JP 1987190083-A/1
PD 06-AUG-1987
PF 31-JAN-1986 JP 1986031415
PI OKAI HIDEO, KUMAKURA TAKESHI, KAWAMOTO SHOICHI, KOIDE TAKAO,
MOMOTA YUTAKA
PC C12N15/00,C07H21/04,C12N1/00,C12P21/02,(C12N1/00,C12R1:19), PC
(C12P21/02);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No; Location/Qualifiers
FH Key
FH Promoter 125..170
FT RBS 200..203
FT sig_peptide 209..277
FT mat_peptide 278..436
FT /product='human beta-urogastrone' FT CDS
209..439
FT /product='human beta-urogastrone'.
FEATURES
source
1..456
/organism='synthetic construct'
/mol_type='genomic DNA'
/db_xref='taxon:32630'
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20; Mismatches 0; Indels 0; Gaps 0;
Matches 101; Conservative 0;
QY 1 AGGGTTATTGTCATGCGGATACATATTTGAATGTTAGAAAAATAAACAATAG 60
|||||
Db 173 AGGGTTATTGTCATGCGGATACATATTTGAATGTTAGAAAAATAAACAATAG 114
|||||
QY 61 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 101
|||||
Db 113 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 73
|||||
RESULT 13
E01302/c
LOCUS
DEFINITION
E01302
ACCESSION
E01302.1 GI:16509129
KEYWORDS
Drosophila melanogaster (fruit fly)
SOURCE
Drosophila melanogaster
ORGANISM
Drosophila melanogaster
REFERENCE
1
AUTHORS
Deak,P., Glover,D.M. and Midgley,C.
TITLE
Cell cycle progression proteins
JOURNAL
Patent: WO 0172774-A 60 04-OCT-2001;
Cyclacel Limited (GB)
FEATURES
source
1..466
/organism='Drosophila melanogaster'
/mol_type='unassigned DNA'
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20; Mismatches 0; Indels 0; Gaps 0;
Matches 101; Conservative 0;
QY 1 AGGGTTATTGTCATGCGGATACATATTTGAATGTTAGAAAAATAAACAATAG 60
|||||
Db 173 AGGGTTATTGTCATGCGGATACATATTTGAATGTTAGAAAAATAAACAATAG 114
|||||
QY 61 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 101
|||||
Db 113 GGGTTCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 73
|||||
```

```
ORIGIN
/db_xref="taxon:7227"

Query Match      100.0%; Score 101; DB 6; Length 466;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTAGAAAATAAACAAATAG 60
    |||||
Db 280 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTAGAAAATAAACAAATAG 221
    |||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
    |||||
Db 220 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 180
    |||||

RESULT 14
AX260150/c      AX260150      573 bp      DNA      linear      PAT 26-OCT-2001
LOCUS
DEFINITION      Sequence 112 from Patent WO0172774.
ACCESSION      AX260150
VERSION      AX260150.1 GI:16509172
KEYWORDS
SOURCE
ORGANISM      Drosophila melanogaster (fruit fly)
                Drosophila melanogaster
                Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
                Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
                Ephydroidea; Drosophilidae; Drosophila.
REFERENCE      1
AUTHORS      Deak, P., Glover, D.M. and Midgley, C.
TITLE      Cell cycle progression proteins
JOURNAL      Patent: WO 0172774-A 112 04-OCT-2001;
                Cyclacel Limited (GB)
FEATURES      Location/Qualifiers
                source
                1..573
                /organism="Drosophila melanogaster"
                /mol_type="unassigned DNA"
                /db_xref="taxon:7227"

ORIGIN

Query Match      100.0%; Score 101; DB 6; Length 573;
Best Local Similarity 100.0%; Pred. No. 8.8e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTAGAAAATAAACAAATAG 60
    |||||
Db 355 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTAGAAAATAAACAAATAG 296
    |||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
    |||||
Db 295 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 255
    |||||

RESULT 15
A43586
LOCUS      A43586      693 bp      DNA      linear      PAT 06-MAR-1997
DEFINITION      Sequence 11 from Patent WO9507357.
ACCESSION      A43586
VERSION      A43586.1 GI:2298779
KEYWORDS
SOURCE      Cuphea lanceolata
ORGANISM      Cuphea lanceolata
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                rosids; Myrtales; Lythraceae; Cuphea.
REFERENCE      1 (bases 1 to 693)
AUTHORS      Toepfer, R., Bautor, J., Bothmann, H., Filsak, E.,
                Hoerhricke-Grandpierre, C., Klein, B., Martini, N., Mueller, A.,
                Schulte, W., Voeltz, M., Walek, J. and Schell, J.
                PROMOTERS
TITLE      Patent: WO 9507357-A 11 16-MAR-1995;
                MAX PLANCK GESELLSCHAFT (DE)
JOURNAL
COMMENT      Other publication CA 2169093 950316
```

```
FEATURES      source
                Location/Qualifiers
                1..693
                /organism="Cuphea lanceolata"
                /mol_type="unassigned DNA"
                /db_xref="taxon:3930"
                /clone="CLKASIG8"
                /clone_lib="Genomic Lambda Fix II"

ORIGIN

Query Match      100.0%; Score 101; DB 6; Length 693;
Best Local Similarity 100.0%; Pred. No. 8.8e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTAGAAAATAAACAAATAG 60
    |||||
Db 592 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTAGAAAATAAACAAATAG 651
    |||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
    |||||
Db 652 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 692
    |||||

Search completed: July 14, 2005, 14:03:32
Job time : 756.618 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:35:42 ; Search time 142.398 Seconds
(without alignments)
4198.742 Million cell updates/sec

Title: US-09-482-682-47_COPY_7889_7989

Perfect score: 101

Sequence: 1 agggatttctctatgacg.....gaaagtgccacctgacgtc 101

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04.*

1: Geneseqn1980s.*

2: Geneseqn1990s.*

3: Geneseqn2000s.*

4: Geneseqn2001as.*

5: Geneseqn2001bs.*

6: Geneseqn2002as.*

7: Geneseqn2002bs.*

8: Geneseqn2003as.*

9: Geneseqn2003bs.*

10: Geneseqn2003cs.*

11: Geneseqn2003ds.*

12: Geneseqn2004as.*

13: Geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	101	100.0	142	2 AAV76919	AAV76919 Staphylococcus aureus contig SEQ ID #2608.
C 2	101	100.0	228	1 AAN10032	Aan10032 Sequence
C 3	101	100.0	251	1 AAN10031	Aan10031 Sequence
C 4	101	100.0	400	2 AAV31229	AAV31229 E. coli J
C 5	101	100.0	456	1 AAN60624	Aan60624 Plasmid p
C 6	101	100.0	456	1 AAN71080	Aan71080 Sequence
C 7	101	100.0	456	1 AAN70833	Aan70833 Beta-urog
C 8	101	100.0	456	1 AAN81765	Aan81765 Sequence
C 9	101	100.0	466	6 ABA90413	Abag90413 Drosophila
C 10	101	100.0	487	2 AAX21173	Aax21173 Polynucle
C 11	101	100.0	535	2 AAX21149	Aax21149 Polynucle
C 12	101	100.0	573	6 ABA90456	Abag90456 Drosophila
C 13	101	100.0	605	12 ADH58311	Adh58311 Electroph
C 14	101	100.0	776	4 AAS30560	Aas30560 DNA encod
C 15	101	100.0	776	4 AAS27819	Aas27819 DNA encod
C 16	101	100.0	776	4 ABK42984	Abk42984 Genomic s
C 17	101	100.0	776	4 AAL07344	Aal07344 Human rep
C 18	101	100.0	776	4 AAL03229	Aal03229 Human rep
C 19	101	100.0	776	4 AAL06588	Aal06588 Human rep
C 20	101	100.0	776	4 AAL07340	Aal07340 Human rep

C 21	101	100.0	776	5 ABA14573	Abal14573 Human ner
C 22	101	100.0	776	5 AAS34681	Aas34681 Human DNA
C 23	101	100.0	776	8 ADA41574	Ada41574 Human sec
C 24	101	100.0	776	8 ACC50905	Acc50905 Human sec
C 25	101	100.0	776	8 ABZ71508	Abz71508 Secreted
C 26	101	100.0	776	9 ADB91869	Adb91869 Human sec
C 27	101	100.0	776	9 ADB61140	Adb61140 Connectiv
C 28	101	100.0	776	10 ADB94622	Adb94622 Novel hum
C 29	101	100.0	776	10 ADC74663	Adc74663 Human sec
C 30	101	100.0	776	10 ADA57709	Ada57709 BAC fragm
C 31	101	100.0	776	12 ADN41551	Adn41551 Novel hum
C 32	101	100.0	845	4 AAS30559	Aas30559 DNA encod
C 33	101	100.0	845	4 AAS27818	Aas27818 DNA encod
C 34	101	100.0	845	4 ABK42983	Abk42983 Genomic s
C 35	101	100.0	845	4 AAS41807	Aas41807 Genomic s
C 36	101	100.0	845	4 AAS41855	Aas41855 Genomic s
C 37	101	100.0	845	4 AAK85485	Aak85485 Human imm
C 38	101	100.0	845	4 AAK85434	Aak85434 Human imm
C 39	101	100.0	845	4 AAL07343	Aal07343 Human rep
C 40	101	100.0	845	4 AAL06587	Aal06587 Human rep
C 41	101	100.0	845	4 AAL07339	Aal07339 Human rep
C 42	101	100.0	845	4 AAL03228	Aal03228 Human rep
C 43	101	100.0	845	5 ABA14572	Abal14572 Human ner
C 44	101	100.0	845	5 AAS34680	Aas34680 Human DNA
C 45	101	100.0	845	9 ADB61139	Adb61139 Connectiv

ALIGNMENTS

RESULT 1

AAV76919/c

ID AAV76919 standard; DNA; 142 BP.

XX AC AAV76919;

XX DT 16-MAR-1999 (first entry)

XX DB Staphylococcus aureus contig SEQ ID #2608.

XX KW Computer readable medium; vaccine; S.aureus infection; immunodetection; cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy; skin infection; surgical wound infection; scalded skin syndrome; toxic shock syndrome; ds.

XX OS Staphylococcus aureus.

XX PN EP786519-A2.

XX PD 30-JUL-1997.

XX PF 07-JAN-1997; 97EP-00100117.

XX PR 05-JAN-1996; 96US-0009861P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Kunsch CA, Choi GH, Barash SC, Dillon PJ, Fannon MR, Rosen CA;

XX DR WPI; 1997-374922/35.

XX PT Polynucleotide(s) and proteins derived from Staphylococcus aureus - stored on computer readable medium and used in the production of anti-S.aureus vaccines.

XX PS Claim 1; Page 2287; 3271pp; English.

XX CC This sequence represents one of 5191 Staphylococcus aureus DNA sequences of the invention. The DNA sequences are recorded on a computer readable medium, preferably selected from a floppy or hard disk, random access memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using the S.aureus DNA sequences allows putative functions to be assigned so that protein-encoding or regulatory regions of commercial, therapeutic or

CC industrial importance can be obtained. Specifically, sequences which are
CC likely to encode antigens have been identified and these polypeptides can
CC be used in a vaccine composition against S.aureus infection. The
CC polypeptides can also be used in a kit for the immunodetection of
CC S.aureus in a sample. S.aureus is implicated in numerous human diseases,
CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,
CC skin and surgical wound infections, scalded skin syndrome, toxic shock
CC syndrome, etc. Organisms transformed with the DNA sequences can be used
CC for recombinant production of the polypeptides. The new DNA sequences
CC (and their fragments) are useful as primers or probes for isolating
CC homologues of any of the S.aureus DNA sequences contained on the computer
CC readable medium.

SQ Sequence 142 BP; 45 A; 25 C; 26 G; 45 T; 0 U; 1 Other;

Query Match 100.0%; Score 101; DB 2; Length 142;

Best Local Similarity 100.0%; Pred. No. 2.1e-21;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGTTATTGTCATGAGCGGATACATATTTGAATGTTATTTAGAAAAATAACAATAG 60

Db 107 AGGTTATTGTCATGAGCGGATACATATTTGAATGTTATTTAGAAAAATAACAATAG 48

Qy 61 GGGTTCCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 101

Db 47 GGGTTCCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 7

RESULT 2

AAN10032/c

ID AAN10032 standard; DNA; 228 BP.

XX AAN10032;

XX 13-AUG-1992 (first entry)

XX Sequence of the pKT218 EcoRI-PstI penicillinase gene fragment.

XX Cloning vehicle; bacterial vector; transformed host; penicillinase;

KW insulin; ds.

XX Escherichia coli.

XX Key Location/Qualifiers

FT misc_feature 1..4

FT /*tag= a

FT /label= sticky end

FT 225..228

FT /*tag= b

FT /label= sticky end

XX EP38182-A.

XX 21-OCT-1981.

XX 09-APR-1981; 81EP-00301561.

XX 11-APR-1980; 80US-00139225.

XX (HARD) HARVARD COLLEGE.

XX Gilbert W, Talmadge K;

XX WPI; 1981-80125D/44.

XX P-PSDB; AAP10039.

XX Synthesis of mature protein or polypeptide - by using bacterial host

XX transformed by cloned vehicle contg. DNA fragment etc.

XX Example; Fig 3; 34pp; English.

XX The closest identifiable promoter for the penicillinase gene in pKT241

CC (AAN10031) is located in the region 14 to 20 nucleotides before its

CC translational start signal. In the examples, the 3' end of pKT241 was
CC attached to the signal DNA sequence of the DNA fragment (19) for rat
CC preproinsulin (see AAN10033). The closest identifiable promoter for the
CC penicillinase gene in pKT218 (AAN10032) is located in the region 14 to 20
CC nucleotides before its translational start signal. In the examples, the
CC 3' end of pKT218 was attached to the signal DNA sequence of the DNA
CC fragment (CB6) for rat preproinsulin (see AAN10034)

SQ Sequence 228 BP; 72 A; 37 C; 46 G; 73 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 1; Length 228;

Best Local Similarity 100.0%; Pred. No. 2.3e-21;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGTTATTGTCATGAGCGGATACATATTTGAATGTTATTTAGAAAAATAACAATAG 60

Db 175 AGGTTATTGTCATGAGCGGATACATATTTGAATGTTATTTAGAAAAATAACAATAG 116

Qy 61 GGGTTCCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 101

Db 115 GGGTTCCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 75

RESULT 3

AAN10031/c

ID AAN10031 standard; DNA; 251 BP.

XX AAN10031;

XX 13-AUG-1992 (first entry)

XX Sequence of the pKT241 EcoRI-PstI penicillinase gene fragment.

XX Cloning vehicle; bacterial vector; transformed host; penicillinase;

KW insulin; ds.

XX Escherichia coli.

XX Key Location/Qualifiers

FT misc_feature 1..4

FT /*tag= a

FT /label= sticky end

FT 248..251

FT /*tag= b

FT /label= sticky end

XX EP38182-A.

XX 21-OCT-1981.

XX 09-APR-1981; 81EP-00301561.

XX 11-APR-1980; 80US-00139225.

XX (HARD) HARVARD COLLEGE.

XX Gilbert W, Talmadge K;

XX WPI; 1981-80125D/44.

XX P-PSDB; AAP10038.

XX Synthesis of mature protein or polypeptide - by using bacterial host

XX transformed by cloned vehicle contg. DNA fragment etc.

XX Example; Fig 2; 34pp; English.

XX The closest identifiable promoter for the penicillinase gene in pKT241

CC (AAN10031) is located in the region 14 to 20 nucleotides before its

CC translational start signal. In the examples, the 3' end of pKT241 was

CC attached to the signal DNA sequence of the DNA fragment (19) for rat

CC preproinsulin (see AAN10033). The closest identifiable promoter for the

CC penicillinase gene in pKT218 (AAN10032) is located in the region 14 to 20

CC nucleotides before its translational start signal. In the examples, the

CC 3' end of pKT218 was attached to the signal DNA sequence of the DNA
CC fragment (CB6) for rat preproinsulin (see AAN10034)

XX Sequence 251 BP; 74 A; 45 C; 50 G; 82 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 1; Length 251;
Best Local Similarity 100.0%; Pred. No. 2.3e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAACAAATAG 60
Db 175 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAACAAATAG 116
QY 61 GGGTTCCGCGCACATTTCCCGAAAAAGTGCCACCTGAGCTC 101
Db 115 GGGTTCCGCGCACATTTCCCGAAAAAGTGCCACCTGAGCTC 75

RESULT 4

AAV31229/c
ID AAV31229 standard; DNA; 400 BP.

XX AC AAV31229;

XX DT 01-OCT-1998 (first entry)

XX DE E. coli J96 pathogenicity island contig #43.

XX KW PAI; pathogenicity island; uropathogenic E. coli detection; PAI IV; pheR;
XX KW PAI V; pheV; vaccine; protective immune response; ds.

XX OS Escherichia coli.

XX PN WO9822575-A2.

XX PD 28-MAY-1998.

XX PF 21-NOV-1997; 97WO-US021347.

XX PR 22-NOV-1996; 96US-0031626P.

XX PR 14-OCT-1997; 97US-0061953P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PA (UYWI-) UNIV WISCONSIN.

XX PI Dillon PJ, Choi GH, Welch RA;

XX WPI; 1998-312461/27.

XX New isolated uropathogenic E. coli nucleotide sequences - used to develop
XX products for the detection of pathogenic E. coli and to elicit an immune
XX response to pathogenic E. coli.

XX Claim 21; Page 140-141; 250pp; English.

XX This sequence represents a E. coli strain J96 contig containing
XX pathogenicity island (PAI) sequences, and represents a nucleic acid
XX molecule of the invention. PAIs are large fragments of DNA which comprise
XX pathogenicity determinants. The sequences of the invention are taken from
XX on the E. coli chromosome and is greater than 170 kb. PAI V is located at
XX approximately 94 min (at pheR) on the E. coli chromosome and is
XX approximately 160 kb in size. Antibodies specific to the proteins encoded
XX by the PAI open reading frames of the invention can be used in kits to
XX detect uropathogenic E. coli. The proteins are used in vaccines to elicit
XX a protective immune response in an animal to the uropathogenic E. coli
XX strain J96

XX Sequence 400 BP; 106 A; 77 C; 91 G; 126 T; 0 U; 0 Other;

XX Query Match

XX Best Local Similarity 100.0%; Score 101; DB 2; Length 400;

XX Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAACAAATAG 60
Db 165 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAACAAATAG 106
QY 61 GGGTTCCGCGCACATTTCCCGAAAAAGTGCCACCTGAGCTC 101
Db 105 GGGTTCCGCGCACATTTCCCGAAAAAGTGCCACCTGAGCTC 65

RESULT 5

AAN60624/c
ID AAN60624 standard; DNA; 456 BP.

XX AC AAN60624;

XX DT 25-MAR-2003 (revised)

XX DT 29-OCT-1991 (first entry)

XX DE Plasmid pUG201 sequence encoding beta-urogastrone.

XX KW Beta-lactamase signal peptide; pGH54; pGH55; ss.

XX OS Synthetic.

XX FH Key Location/Qualifiers

XX FT promoter 125..170

XX FT /*tag= a

XX FT RBS 200..203

XX FT /*tag= b

XX FT CDS 209..439

XX FT /*tag= c

XX FT sig_peptide 209..277

XX FT /*tag= d

XX FT /label= Beta-lactamase signal peptide

XX FT 278..436

XX FT /*tag= e

XX FT /label= Beta-urogastrone

XX W08603779-A.

XX 03-JUL-1986.

XX PF 19-DEC-1985; 85WO-JP000696.

XX 21-DEC-1984; 84JP-00271206.

XX (EART) EARTH CHEM CO LTD.

XX (OHGA/) OHGAI H.

XX Ohgai H, Momota H, Kuwakura T, Kajifusa N, Kitazawa T, Oshiden K;

XX WPI; 1986-182911/28.

XX P-PSDB; AAP60678.

XX Recombinant vector for polypeptide secretion - contains signal peptide

XX sequence directly bonded to peptide-coding sequence.

XX Disclosure; Table 4; 79pp; Japanese.

XX The plasmid produces secreted beta-urogastrone in a transformed
XX expression system. Similar plasmids may be constructed where the
XX secretion signal may be coupled with eg. somatostatin, insulin, growth
XX hormone, interferon, IL-2, gastric inhibitory peptide, influenza B SA,
XX epidermal growth factor and thymosine-beta4. (Updated on 25-MAR-2003 to
XX correct PA field.)

XX SQ Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;

XX Query Match

XX Best Local Similarity 100.0%; Score 101; DB 1; Length 456;

XX Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTATTAGAAAATAAACAATAG 60
|||||
Db 173 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTATTAGAAAATAAACAATAG 114
|||||
QY 61 GGGTTCCGCGACATTTCCCGGAAAAGTGCACCTGACGTC 101
|||||
Db 113 GGGTTCCGCGACATTTCCCGGAAAAGTGCACCTGACGTC 73
|||||

RESULT 6

AAAN71080/c
ID AAAN71080 standard; DNA; 456 BP.
XX
AC AAAN71080;
XX
DT 25-MAR-2003 (revised)
DT 10-MAR-2003 (revised)
DT 13-MAY-1991 (first entry)
XX
DE Sequence encoding beta-urogastrone.
XX
KW PUGT 150s; beta-UG; ds.
XX
XX Escherichia coli.
OS Synthetic.
XX
XX
FH Key Location/Qualifiers
FT promoter 125..170
FT /*tag= a
FT CDS 209..439
FT /*tag= b
FT /*transl_except= (pos:434..436,aa:Arg)
XX

JPG2190083-A.

XX
XX 20-AUG-1987.
XX
PF 14-FEB-1986; 86JP-00031415.
XX
PR 14-FEB-1986; 86JP-00031415.
XX
PA (EART) EARTH SEIYAKU KK.
XX
XX
DR WPI; 1987-273761/39.
XX
XX Expression vector contg. multiple information units is used to transform host all for increased prodn. of polypeptides.
PT
PS Disclosure; Page 553; 34pp; Japanese.
XX
XX Sequence encodes beta-urogastrone under the control of a tac promoter. The peptide may be expressed from plasmid pUGT 150s in a transformed E.coli host. The plasmid may carry several separately expressing CC sequences comprising a tac promoter, SD site, signal peptide, and coding CC sequence, to produce beta-UG in high yield. (Updated on 10-MAR-2003 to CC add missing OS field.) (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 1; Length 456;
Best Local Similarity 100.0%; Pred. No. 2.6e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTATTAGAAAATAAACAATAG 60
|||||
Db 173 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTATTAGAAAATAAACAATAG 114
|||||
QY 61 GGGTTCCGCGACATTTCCCGGAAAAGTGCACCTGACGTC 101
|||||
Db 113 GGGTTCCGCGACATTTCCCGGAAAAGTGCACCTGACGTC 73
|||||

RESULT 7

AAAN70833/c
ID AAAN70833 standard; DNA; 456 BP.
XX
AC AAAN70833;
XX
DT 25-MAR-2003 (revised)
DT 10-MAR-2003 (revised)
DT 18-JAN-1991 (first entry)
XX
DE Beta-urogastrone sequence.
XX
XX Tumour; inosine; DNA probe; ds.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT promoter 125..170
FT /*tag= b
FT RBS 200..204
FT /*tag= C
FT CDS 209..439
FT /*tag= a
FT sig_peptide 209..277
FT /*tag= d
XX
XX JP62244398-A.
XX
XX 24-OCT-1987.
XX
XX 16-APR-1986; 86JP-00087368.
XX
XX 16-APR-1986; 86JP-00087368.
XX
XX (SEKI) SEKISUI CHEM IND CO LTD.
XX
XX WPI; 1987-339045/48.
DR P-PSDB; AAP70505.
XX
XX Detection of DNA and/or RNA - by converting to single strand form and using probe contg. labelled inosine deriv.
PT
XX
XX Disclosure; Page 11; 11pp; Japanese.
XX
XX An example of a sequence detected by a probe consisting of polyinosine, polydeoxyinosine, oligoiosine and/or oligodeoxyinosine is labeled. The CC sDNA and probe are hybridized and the existence of DNA in the product is CC detected. It can be used to detect the presence of malignant tumour. CC (Updated on 10-MAR-2003 to add missing OS field.) (Updated on 25-MAR-2003 CC to correct PA field.)
XX
SQ Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 1; Length 456;
Best Local Similarity 100.0%; Pred. No. 2.6e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTATTAGAAAATAAACAATAG 60
|||||
Db 173 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTTATTAGAAAATAAACAATAG 114
|||||
QY 61 GGGTTCCGCGACATTTCCCGGAAAAGTGCACCTGACGTC 101
|||||
Db 113 GGGTTCCGCGACATTTCCCGGAAAAGTGCACCTGACGTC 73
|||||

RESULT 8

AAAN81765/c
ID AAAN81765 standard; DNA; 456 BP.
XX
XX AAAN81765;
XX
XX
DT 25-MAR-2003 (revised)
DT 13-DEC-1990 (first entry)

XX Sequence of HB101 (pGH201) encoding new beta-urogastrone deriv. Met (21),
DE Arg (53).
XX Gastric acid secretion; cell proliferation; hormone; ds.
XX Synthetic.
XX Key Location/Qualifiers
FH 209. .277
FT CDS /*tag= a
FT 278. .439
FT CDS /*tag= b
FT /product= "New beta-urogastrone deriv."
XX JP63012298-A.
XX 19-JAN-1988.
XX 30-JUN-1986; 86JP-00153783.
XX 30-JUN-1986; 86JP-00153783.
XX (EART) EARTH SEIYAKU KK.
XX WPI; 1988-054638/08.
XX P-PSDB; AAP81349.
XX New beta-urogastrone deriv. - has gastric acid secretion inhibition and
PT proliferation promotion activity.
XX Disclosure; Page 685; 76pp; Japanese.
XX The deriv. has various biological activities such as gastric acid
CC secretion inhibiting action, or cell proliferation promoting action. The
CC deriv. has the same biological or pharmacological activities as beta-
CC urogastrone. It is not susceptible to denaturation by oxidn. and is
CC chemically stable. Deriv. has resistance to proteolytic enzymes such as
CC protease. (Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 1; Length 456;
Best Local Similarity 100.0%; Pred. No. 2.6e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAATAG 60
DB 173 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAATAG 114
QY 61 GGGTTCCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101
DB 113 GGGTTCCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 73
RESULT 9
ABA90413/C
ID ABA90413 standard; DNA; 466 BP.
XX
XX ABA90413;
XX
XX 12-FEB-2002 (first entry)
XX Drosophila cell cycle progression protein coding sequence #48.
DE Antiproliferative; cytostatic; cardiant; immunosuppressive; meiosis;
XX antifungal; antifungal; antifungal; antifungal; antifungal;
XX antiparasitic; antiparasitic; antiparasitic; antiparasitic; antiparasitic;
XX cell cycle progression protein; tumour; proliferative disorder;
XX cardiovascular; autoimmune; dermatological disorder; ds.
XX Drosophila sp.
XX

PN WO200172774-A2.
XX 04-OCT-2001.
XX 23-MAR-2001; 2001WO-GB001297.
XX 24-MAR-2000; 2000GB-00007268.
XX (CYCL-) CYCLACEL LTD.
XX Deak P, Glover DM, Midgley C;
XX WPI; 2002-055132/07.
XX Polynucleotides encoding cell cycle progression proteins, useful for
PT treating a tumor or a proliferative disorder.
XX Claim 1; Page 99; 213pp; English.
XX The present invention relates to Drosophila cell cycle progression
CC proteins (AAM47572-AAM47608) and their coding sequences (ABA90366-
CC ABA90520). The coding sequences and proteins are useful for identifying a
CC substance capable of affecting the function of the corresponding gene, a
CC substance capable of inhibiting the cell division cycle, or capable of
CC inhibiting mitosis and/or meiosis. They can also be used in a method for
CC treating a tumour or proliferative disorder, cardiovascular disorders
CC (such as restenosis and cardiomyopathy), autoimmune disorders such as
CC (glomerulonephritis and rheumatoid arthritis), dermatological disorders
CC (such as psoriasis), antiinflammatory, antifungal and antiparasitic
CC disorders (such as malaria)
XX
XX Sequence 466 BP; 135 A; 87 C; 93 G; 150 T; 0 U; 1 Other;
Query Match 100.0%; Score 101; DB 6; Length 466;
Best Local Similarity 100.0%; Pred. No. 2.6e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAATAG 60
DB 280 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAATAG 221
QY 61 GGGTTCCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101
DB 220 GGGTTCCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 180
RESULT 10
AAX211173/C
ID AAX211173 standard; DNA; 487 BP.
XX
XX AAX211173;
XX
XX 05-MAY-1999 (first entry)
XX Polynucleotide sequence from the genome of Treponema pallidum.
DE Treponema pallidum infection; syphilis; Borrelia infection; animal;
XX enzyme production; ds.
XX Treponema pallidum.
XX WO9859034-A2.
XX 30-DEC-1998.
XX 23-JUN-1998; 98WO-US013041.
XX 24-JUN-1997; 97US-0050667P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Fraser CM;
XX

PR	06-SEP-2000;	2000US-02304338P
PR	08-SEP-2000;	2000US-02312422P
PR	08-SEP-2000;	2000US-02312423P
PR	08-SEP-2000;	2000US-02312433P
PR	08-SEP-2000;	2000US-02312443P
PR	08-SEP-2000;	2000US-02314133P
PR	08-SEP-2000;	2000US-02314143P
PR	08-SEP-2000;	2000US-02320808P
PR	08-SEP-2000;	2000US-02320818P
PR	12-SEP-2000;	2000US-02319698P
PR	14-SEP-2000;	2000US-02323978P
PR	14-SEP-2000;	2000US-02323988P
PR	14-SEP-2000;	2000US-02323998P
PR	14-SEP-2000;	2000US-02323999P
PR	14-SEP-2000;	2000US-02324008P
PR	14-SEP-2000;	2000US-02324018P
PR	14-SEP-2000;	2000US-02324019P
PR	14-SEP-2000;	2000US-02324033P
PR	14-SEP-2000;	2000US-02323064P
PR	14-SEP-2000;	2000US-02323065P
PR	21-SEP-2000;	2000US-02342233P
PR	21-SEP-2000;	2000US-02342274P
PR	21-SEP-2000;	2000US-02342747P
PR	21-SEP-2000;	2000US-02349978P
PR	25-SEP-2000;	2000US-02349998P
PR	26-SEP-2000;	2000US-02345848P
PR	27-SEP-2000;	2000US-02358348P
PR	27-SEP-2000;	2000US-02358349P
PR	27-SEP-2000;	2000US-02358368P
PR	27-SEP-2000;	2000US-02358369P
PR	29-SEP-2000;	2000US-02363676P
PR	29-SEP-2000;	2000US-02363677P
PR	29-SEP-2000;	2000US-02370378P
PR	02-OCT-2000;	2000US-02370379P
PR	02-OCT-2000;	2000US-02370393P
PR	02-OCT-2000;	2000US-02370408P
PR	13-OCT-2000;	2000US-02399338P
PR	13-OCT-2000;	2000US-02399339P
PR	20-OCT-2000;	2000US-02409650P
PR	20-OCT-2000;	2000US-02409651P
PR	20-OCT-2000;	2000US-02418098P
PR	20-OCT-2000;	2000US-02418099P
PR	01-NOV-2000;	2000US-02461778P
PR	01-NOV-2000;	2000US-02461785P
PR	20-OCT-2000;	2000US-02417868P
PR	20-OCT-2000;	2000US-02417878P
PR	20-OCT-2000;	2000US-02418088P
PR	20-OCT-2000;	2000US-02418098P
PR	20-OCT-2000;	2000US-02418099P
PR	08-NOV-2000;	2000US-02464674P
PR	08-NOV-2000;	2000US-02464745P
PR	08-NOV-2000;	2000US-02464758P
PR	08-NOV-2000;	2000US-02464768P
PR	08-NOV-2000;	2000US-02464778P
PR	08-NOV-2000;	2000US-02464788P
PR	08-NOV-2000;	2000US-02465238P
PR	08-NOV-2000;	2000US-02465248P
PR	08-NOV-2000;	2000US-02465258P
PR	08-NOV-2000;	2000US-02465268P
PR	08-NOV-2000;	2000US-02465278P
PR	08-NOV-2000;	2000US-02465288P
PR	08-NOV-2000;	2000US-02465298P
PR	08-NOV-2000;	2000US-02465308P
PR	08-NOV-2000;	2000US-02465318P
PR	08-NOV-2000;	2000US-02465328P
PR	08-NOV-2000;	2000US-02465338P
PR	08-NOV-2000;	2000US-02465348P
PR	08-NOV-2000;	2000US-02465358P
PR	08-NOV-2000;	2000US-02465368P
PR	08-NOV-2000;	2000US-02465378P
PR	08-NOV-2000;	2000US-02465388P
PR	08-NOV-2000;	2000US-02465398P
PR	08-NOV-2000;	2000US-02465408P
PR	08-NOV-2000;	2000US-02465418P
PR	08-NOV-2000;	2000US-02465428P
PR	08-NOV-2000;	2000US-02465438P
PR	08-NOV-2000;	2000US-02465448P
PR	08-NOV-2000;	2000US-02465458P
PR	08-NOV-2000;	2000US-02465468P
PR	08-NOV-2000;	2000US-02465478P
PR	08-NOV-2000;	2000US-02465488P
PR	08-NOV-2000;	2000US-02465498P
PR	08-NOV-2000;	2000US-02465508P
PR	08-NOV-2000;	2000US-02465518P
PR	08-NOV-2000;	2000US-02465528P
PR	08-NOV-2000;	2000US-02465538P
PR	08-NOV-2000;	2000US-02465548P
PR	08-NOV-2000;	2000US-02465558P
PR	08-NOV-2000;	2000US-02465568P
PR	08-NOV-2000;	2000US-02465578P
PR	08-NOV-2000;	2000US-02465588P
PR	08-NOV-2000;	2000US-02465598P

PR	17-NOV-2000;	2000US-0249244P.
PR	17-NOV-2000;	2000US-0249245P.
PR	17-NOV-2000;	2000US-0249246P.
PR	17-NOV-2000;	2000US-0249247P.
PR	17-NOV-2000;	2000US-0249265P.
PR	17-NOV-2000;	2000US-0249266P.
PR	17-NOV-2000;	2000US-0249267P.
PR	17-NOV-2000;	2000US-0249299P.
PR	17-NOV-2000;	2000US-0249300P.
PR	01-DEC-2000;	2000US-0250160P.
PR	01-DEC-2000;	2000US-0250319P.
PR	05-DEC-2000;	2000US-0251030P.
PR	05-DEC-2000;	2000US-0251588P.
PR	05-DEC-2000;	2000US-0256719P.
PR	06-DEC-2000;	2000US-0251479P.
PR	08-DEC-2000;	2000US-0251856P.
PR	08-DEC-2000;	2000US-0251868P.
PR	08-DEC-2000;	2000US-0251869P.
PR	08-DEC-2000;	2000US-0251989P.
PR	08-DEC-2000;	2000US-0251990P.
PR	11-DEC-2000;	2000US-0254097P.
PR	05-JAN-2001;	2000US-0259678P.
XX	(HUMA-)	HUMAN GENOME SCI INC.
PA		

Novel isolated prostate gland related polypeptide useful for diagnosis and treatment of disorders of prostate such as prostatodystonia, prostatosis, prostatitis, benign prostatic hypertrophy and malacoplakia.

Claim 1: SEQ ID NO 418: 512bp; English.

The invention relates to novel isolated prostate gland related nucleic acids (I) and polypeptides (II). (I) and (II) are useful for diagnosis, prognosis, prevention, and/or treatment of diseases and/or disorders of the prostate such as acute non-bacterial prostatitis, chronic non-bacterial prostatitis, acute bacterial prostatitis, prostatic dysplasia, prostatic adenocarcinoma, granulomatous prostatitis, malacoplakia, benign prostatic hyperplasia, and prostate neoplastic disorders, including adenocarcinomas, transitional cell carcinomas, ductal carcinomas, and squamous cell carcinomas. (I), (II) and antibody to (II) are useful for diagnosing and treating reproductive system disorders (Paget's disease), autoimmune disorders (systemic lupus erythematosus, rheumatoid arthritis), blood-related disorders (sickle cell anaemia), hyperproliferative disorders, urinary system disorders (glomerulonephritis), cardiovascular disorders (arrhythmias), respiratory disorders, musculoskeletal system disorders, neural activity and neurological disorders (Alzheimer's disease and Parkinson's disease), endocrine disorders (Addison's disease), gastrointestinal disorders (inflammatory disorders), liver disorders (biliary liver cirrhosis), pancreatic and gall bladder disorders, disorders of the large intestine, developmental and inherited disorders, diseases at the cellular level, and wound healing and epithelial cell proliferation. (I) or (II) is useful to prevent skin aging, for preventing hair loss, to maintain organs before transplantation, and as food additive or preservative.

Query Match	100.0%	Score 101;	DB 4;	Length 776;
Best Local Similarity	100.0%	Pred. No. 2.9e-21;		
Matches 101;	Conservative	0;	Mismatches 0;	Indels 0;
				Gaps 0;

QY	1	AGGGTTATTGTTCTCATGAGCGGATACATATTTGAAATGTTATTTAGAAAAATATAACAAATAG	60
Db	546	AGGGTTATTGTTCTCATGAGCGGATACATATTTGAAATGTTATTTAGAAAAATATAACAAATAG	487
QY	61	GGGTTCCCGCACATTTCCCGAAAAAGTGCCACCTGACGTC	101
Db	486	GGGTTCCCGCACATTTCCCGAAAAAGTGCCACCTGACGTC	446

RESULT 15
AAS27819/C
ID AAS27819 standard; DNA: 776 BP.

XX AAS27819;
AC
XX
DT 07-NOV-2001 (first entry)
XX
XX DNA encoding novel signal transduction pathway protein, Seq ID 1479.
XX
KW Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
KW antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;
KW immune system disorder; rheumatoid arthritis; inflammatory condition;
KW organ transplant rejection; infection; hepatitis C; blood disorder;
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
KW reproductive system; gastrointestinal; liver disorder; AIDS; ds;
KW acquired immune deficiency syndrome.
XX
OS Homo sapiens.
XX
XX WO200154733-A1.
XX
XX PD 02-AUG-2001.
XX
XX PF 17-JAN-2001; 2001WO-US001312.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205151P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 25-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 26-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249265P.

PR	17-NOV-2000;	2000US-0249297P.	
PR	17-NOV-2000;	2000US-0249299P.	
PR	17-NOV-2000;	2000US-0249300P.	
PR	01-DEC-2000;	2000US-0250160P.	
PR	01-DEC-2000;	2000US-0250391P.	
PR	05-DEC-2000;	2000US-0251030P.	
PR	05-DEC-2000;	2000US-0251988P.	
PR	05-DEC-2000;	2000US-0256719P.	
PR	06-DEC-2000;	2000US-0251479P.	
PR	08-DEC-2000;	2000US-0251856P.	
PR	08-DEC-2000;	2000US-0251868P.	
PR	08-DEC-2000;	2000US-0251869P.	
PR	08-DEC-2000;	2000US-0251989P.	
PR	08-DEC-2000;	2000US-0251990P.	
PR	11-DEC-2000;	2000US-0254097P.	
PR	05-JAN-2001;	2001US-0259678P.	
XX			
PA	(HUMA-)	HUMAN GENOME SCI INC.	
XX			
PI	Rosen CA, Barash SC, Ruben SM;		
XX			
DR	WPI; 2001-465460/50.		
XX			
XX			
PT	Novel polypeptides useful for diagnosing, treating, preventing and/or		
PT	prognosing disorders related to the proteins, including cancers, immune		
PT	disorders and neuronal disorders.		
XX			
PS	Claim 1; SEQ ID NO 1479; 880pp; English.		
XX			
CC	The invention relates to novel isolated polypeptides (I), and		
CC	polynucleotides (II). (I) and the antibody to (I) are useful for		
CC	diagnosing, preventing and treating diseases including immune system		
CC	disorders (e.g. congenital and acquired immunodeficiencies, autoimmune		
CC	disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ		
CC	transplant rejections and graft versus host disease, infectious diseases		
CC	(e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and		
CC	other blood-related disorders (sickle cell anaemia), myeloproliferative		
CC	disorders, primary haematopoietic disorders, hyperproliferative disorders		
CC	(e.g. Gaucher's disease and cancer), neurodegenerative disorders (e.g.		
CC	Alzheimer's disease, Parkinson's disease), chromosomal abnormalities		
CC	(Down syndrome), ischaemic injury (e.g. stroke), renal disorders (e.g.		
CC	glomerulonephritis), cardiovascular disorders (e.g. arrhythmia),		
CC	respiratory disorders, dermatological disorders, in wound healing,		
CC	epithelial cell proliferation, endocrine disorders (e.g. Addison's		
CC	disease), reproductive system disorders, gastrointestinal disorder		
CC	(inflammatory disorders), liver disorders (cirrhosis), as stimulators of		
CC	B-cell responsiveness to pathogens, activators of T-cells, to induce		
CC	higher affinity antibodies, and as a means to induce tumour proliferation		
CC	in pathologies e.g. acquired immune deficiency syndrome (AIDS). AAS26976-		
CC	AAS27850 represent novel signal transduction pathway protein coding		
CC	sequences and PCR primers of the invention		
XX			
SQ	Sequence 776 BP; 184 A; 209 C; 181 G; 200 T; 0 U; 2 Other;		
	Query Match 100.0%; Score 101; DB 4; Length 776;		
	Best Local Similarity 100.0%; Pred. No. 2.9e-21;		
	Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0		
QY	1 AGGGTTATTGTTCTCATGAGCGGATACATATTTGAATGTATTATTAGAAAAATAACAAATAG 60		
Db	546 AGGGTTATTGTTCTCATGAGCGGATACATATTTGAATGTATTATTAGAAAAATAACAAATAG 487		
QY	61 GGGTTCCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 101		
Db	486 GGGTTCCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 446		

GenCore version 5.1.6

Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 05:15:57 ; Search time 961.667 Seconds
(without alignments)
3997.736 Million cell updates/sec

Title: US-09-482-682-47_COPY_7889_7989

Perfect score: 101

Sequence: 1 aggggtattgtctatgagc.....gaaagtgccacctgacgtc 101

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hic:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gsa1:*

9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	101	100.0	300	4	BM078095
C 2	101	100.0	300	5	BU963956
C 3	101	100.0	300	5	BU964094
C 4	101	100.0	309	9	FR0009140
C 5	101	100.0	391	1	AL5971149
C 6	101	100.0	414	9	CC819240
C 7	101	100.0	417	4	BJ684174
C 8	101	100.0	491	9	CC819233
C 9	101	100.0	495	4	BI805285
C 10	101	100.0	495	9	CC818374
C 11	101	100.0	496	9	CC818523
C 12	101	100.0	503	9	CC819854
C 13	101	100.0	515	9	CC817752
C 14	101	100.0	518	9	CC817128
C 15	101	100.0	519	9	CC817162
C 16	101	100.0	519	9	CC817796
C 17	101	100.0	521	9	CC819067
C 18	101	100.0	533	9	CC819841
C 19	101	100.0	542	9	CC816892
C 20	101	100.0	550	7	CR766622
C 21	101	100.0	551	9	CC816905
C 22	101	100.0	554	9	CC819058
C 23	101	100.0	563	9	CC819270
C 24	101	100.0	566	9	CC816848

C 25	101	100.0	566	9	CC820024
C 26	101	100.0	567	9	CC817070
C 27	101	100.0	568	9	CC818640
C 28	101	100.0	571	9	CC816954
C 29	101	100.0	571	9	CC818423
C 30	101	100.0	580	9	CC819098
C 31	101	100.0	582	1	AL694813
C 32	101	100.0	583	9	CC817633
C 33	101	100.0	583	9	CC818436
C 34	101	100.0	586	9	CC816883
C 35	101	100.0	588	9	CC817788
C 36	101	100.0	588	9	CC818340
C 37	101	100.0	589	9	CC817595
C 38	101	100.0	590	9	CC819754
C 39	101	100.0	592	9	CC817679
C 40	101	100.0	592	9	CC818508
C 41	101	100.0	593	9	CC816942
C 42	101	100.0	593	9	CC817699
C 43	101	100.0	594	9	CC818287
C 44	101	100.0	594	9	CC818422
C 45	101	100.0	595	9	CC816929

ALIGNMENTS

RESULT 1

BM078095/c

LOCUS

DEFINITION

83374 Hebeloma cylindrosporum functional cDNA library Hebeloma

cylindrosporum cDNA 5', mRNA sequence.

ACCESSION

BM078095

VERSION

BM078095.1

KEYWORDS

EST.

SOURCE

Hebeloma cylindrosporum

ORGANISM

Hebeloma cylindrosporum

REFERENCE

1 (bases 1 to 300)

AUTHORS

Wipf, D., Benjidia, M., Tegeder, M. and Frommer, W.B.

TITLE

Construction of a functional cDNA library from the ectomycorrhizal fungus Hebeloma cylindrosporum

JOURNAL

Unpublished (2001)

COMMENT

Contact: Wipf D.

ZMBP - Center for Molecular Biology of Plants

University of Tuebingen

Auf der Morgenstelle 1, 72076 Tuebingen, Germany

Tel: 49 7071 2976160

Fax: 49 7071 293287

Email: daniel.wipf@zmbp.uni-tuebingen.de

PCR Primers

FORWARD: pDR196 5' primer (PMA 5')

High quality sequence stop: 300

POLYA=No.

FEATURES

Location/Qualifiers

1..300

/organism="Hebeloma cylindrosporum"

/molecule="mRNA"

/strain="H1"

/db_xref="taxon:76867"

/tissue_type="Mycelia"

/lab_host="E. coli XLI-Blue"

/clone_lib="Hebeloma cylindrosporum functional cDNA library"

/note=vector: pDR 196 (unpublished); Site_1: EcoRI;

Site_2: XhoI"

ORIGIN

Query Match

Best Local Similarity

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AGGGTTATTGTCATGCGGATACATATTGTAATGTTAGTATTGTTAGAAAATAACAAATAG 60

```

|||||
174 AGGGTTATTGCTCATGCGGATACATATTGATGATTTAGAAAAATAACAATAG 115
|||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGAGTC 101
|||||
Db 114 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGAGTC 74
|||||

RESULT 2
BU963956/c
LOCUS
DEFINITION EST88 Hebeloma cylindrosporum functional cDNA library Hebeloma
cylindrosporum cDNA, mRNA sequence.
ACCESSION BU963956
VERSION BU963956.1 GI:24204753
KEYWORDS EST.
SOURCE Hebeloma cylindrosporum
ORGANISM Hebeloma cylindrosporum
Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
Agaricales; Cortinariaceae; Hebeloma.
REFERENCE
AUTHORS Wipf,D., Benjdia,M., Rikirsch,E., Zimmermann,S., Tegeder,M. and
Frommer,W.B.
TITLE A Hebeloma cylindrosporum expression cDNA library for suppression
cloning in yeast mutants, complementation of a yeast his4 mutant
and EST analysis
JOURNAL Unpublished (2002)
COMMENT ZMBP - Center for Molecular Biology of Plants
University of Tuebingen
Auf der Morgenstelle 1, 72076 Tuebingen, Germany
Tel: 49 7071 2976160
Fax: 49 7071 293287
Email: daniel.wipf@zmbp.uni-tuebingen.de
homolog to Enterobacteria phage f1 ; ampicillinase (1e-10).
FEATURES
source
1...300
/organism="Hebeloma cylindrosporum"
/mol_type="mRNA"
/strain="H1"
/db_xref="taxon:76867"
/tissue_type="Mycelia"
/lab_host="E. coli XL1-Blue"
/clone_lib="Hebeloma cylindrosporum functional cDNA
library"
/notes="Vector: pDR 196 (unpublished); Site_1: EcoRI;
Site_2: XhoI"

ORIGIN
Query Match 100.0%; Score 101; DB 5; Length 300;
Best Local Similarity 100.0%; Pred. No. 8.1e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGATGATTTAGAAAAATAACAATAG 60
|||||
Db 170 AGGGTTATTGCTCATGAGCGGATACATATTGATGATTTAGAAAAATAACAATAG 111
|||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGAGTC 101
|||||
Db 110 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGAGTC 70
|||||

RESULT 4
FR0009140
LOCUS
DEFINITION F.rubripes GSS sequence, clone 010H20aC4, genomic survey sequence.
ACCESSION AL000426
VERSION AL000426.1 GI:2438278
KEYWORDS GSS; genome survey sequence.
SOURCE Takifugu rubripes (Fugu rubripes)
ORGANISM Takifugu rubripes
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontoidea; Tetraodontidae; Takifugu.
REFERENCE
AUTHORS Elgar,G., Clark,M.S., Meek,S., Smith,S., Warner,S., Edwards,Y.J.,
Bouchireb,N., Cottage,A., Yeo,G.S., Umrانيا,Y., Williams,G. and
Brenner,S.
TITLE Generation and analysis of 25 Mb of genomic DNA from the pufferfish
Fugu rubripes by sequence scanning
JOURNAL Genome Res. 9 (10), 960-971 (1999)
MEDLINE 99455097
PUBMED 10523524
REFERENCE
AUTHORS Elgar,G., Clark,M., Smith,S., Meek,S., Warner,S., Umrانيا,Y.,
Williams,G. and Brenner,S.
TITLE Direct Submission
JOURNAL Submitted (09-SEP-1997) MRC Human Genome Mapping Project Resource
Centre Hinxton, Cambridge, CB10 1SB. Email: biohelp@hgmpr.mrc.ac.uk
Vector: pBluescript II KS
COMMENT V_type: phagemid

```

```

Agaricales; Cortinariaceae; Hebeloma.
REFERENCE
AUTHORS Wipf,D., Benjdia,M., Rikirsch,E., Zimmermann,S., Tegeder,M. and
Frommer,W.B.
TITLE A Hebeloma cylindrosporum expression cDNA library for suppression
cloning in yeast mutants, complementation of a yeast his4 mutant
and EST analysis
JOURNAL Unpublished (2002)
COMMENT Contact: Wipf D
ZMBP - Center for Molecular Biology of Plants
University of Tuebingen
Auf der Morgenstelle 1, 72076 Tuebingen, Germany
Tel: 49 7071 2976160
Fax: 49 7071 293287
Email: daniel.wipf@zmbp.uni-tuebingen.de
no homolog below 1e-10.
FEATURES
Location/Qualifiers
source
1...300
/organism="Hebeloma cylindrosporum"
/mol_type="mRNA"
/strain="H1"
/db_xref="taxon:76867"
/tissue_type="Mycelia"
/lab_host="E. coli XL1-Blue"
/clone_lib="Hebeloma cylindrosporum functional cDNA
library"
/notes="Vector: pDR 196 (unpublished); Site_1: EcoRI;
Site_2: XhoI"

ORIGIN
Query Match 100.0%; Score 101; DB 5; Length 300;
Best Local Similarity 100.0%; Pred. No. 8.1e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGATGATTTAGAAAAATAACAATAG 60
|||||
Db 170 AGGGTTATTGCTCATGAGCGGATACATATTGATGATTTAGAAAAATAACAATAG 111
|||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGAGTC 101
|||||
Db 110 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGAGTC 70
|||||

RESULT 4
FR0009140
LOCUS
DEFINITION F.rubripes GSS sequence, clone 010H20aC4, genomic survey sequence.
ACCESSION AL000426
VERSION AL000426.1 GI:2438278
KEYWORDS GSS; genome survey sequence.
SOURCE Takifugu rubripes (Fugu rubripes)
ORGANISM Takifugu rubripes
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontoidea; Tetraodontidae; Takifugu.
REFERENCE
AUTHORS Elgar,G., Clark,M.S., Meek,S., Smith,S., Warner,S., Edwards,Y.J.,
Bouchireb,N., Cottage,A., Yeo,G.S., Umrانيا,Y., Williams,G. and
Brenner,S.
TITLE Generation and analysis of 25 Mb of genomic DNA from the pufferfish
Fugu rubripes by sequence scanning
JOURNAL Genome Res. 9 (10), 960-971 (1999)
MEDLINE 99455097
PUBMED 10523524
REFERENCE
AUTHORS Elgar,G., Clark,M., Smith,S., Meek,S., Warner,S., Umrانيا,Y.,
Williams,G. and Brenner,S.
TITLE Direct Submission
JOURNAL Submitted (09-SEP-1997) MRC Human Genome Mapping Project Resource
Centre Hinxton, Cambridge, CB10 1SB. Email: biohelp@hgmpr.mrc.ac.uk
Vector: pBluescript II KS
COMMENT V_type: phagemid

```

```

PRIMER: KS
DESCR:
One pass dye-terminator sequencing of cosmid cloned genomic
sequence.
FEATURES
    source
        Location/Qualifiers
            1..309
                /organism="Takifugu rubripes"
                /mol_type="genomic DNA"
                /db_xref="taxon:31033"
                /clone="010H20aC4"
                /clone_lib="cosmid 010H20"
ORIGIN
Query Match      100.0%; Score 101; DB 9; Length 309;
Best Local Similarity 100.0%; Pred. No. 8.2e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAACAATAG 60
Db 39 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAACAATAG 98
Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
Db 99 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 139
RESULT 5
AL597149
LOCUS
DEFINITION DKFZp313J1611_r1 313 (synonym: hicc2) Homo sapiens cDNA clone
AL597149
ACCESSION
VERSION
KEYWORDS
SOURCE EST.
ORGANISM Homo sapiens (human)
REFERENCE
    1 (bases 1 to 391)
AUTHORS Koehrer,K., Beyer,A., Mewes,W., Weil,B. and Wiemann,S.
TITLE EST (Koehrer,K., Beyer,A., Mewes,H.W., Weil,B. and Wiemann,S.)
JOURNAL Unpublished (1999)
COMMENT Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by BPFZ (Biomedical Research Center at the Charite,
Berlin/Germany) within the cDNA sequencing consortium of the German
Genome Project.
No sl sequence available.
This clone (DKFZp313J1611) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
    source
        Location/Qualifiers
            1..391
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="DKFZp313J1611"
                /dev_stages="adult"
                /lab_host="DH10B"
                /clone_lib="313 (synonym: hicc2)"
                /note="vector: priplex2; Site_1: sf1A; Site_2: sf1B;
                cDNA-collection"
ORIGIN
Query Match      100.0%; Score 101; DB 1; Length 391;
Best Local Similarity 100.0%; Pred. No. 8.3e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAACAATAG 60

```

```

Db 228 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAACAATAG 287
Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
Db 288 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 328
RESULT 6
CC819240/c
LOCUS
DEFINITION CC819240 414 bp DNA linear GSS 17-JUL-2003
histriomuscorum genomic clone UUGC10005D19 R, genomic survey
sequence.
ACCESSION CC819240
VERSION CC819240.1 GI:32899308
KEYWORDS GSS.
SOURCE Sterkiella histriomuscorum (Oxytricha trifallax)
ORGANISM Sterkiella histriomuscorum
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
Stichotrichida; Oxytrichidae; Sterkiella.
1 (bases 1 to 414)
Dunn,D., Doak,T., Herrick,G. and Weiss,R.
Paired end reads from plasmid inserts of Oxytricha trifallax
macronuclear chromosomes
Unpublished (2003)
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Plate: 0005 row: D column: 19
Seq primer: CACACGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 414.
FEATURES
    source
        Location/Qualifiers
            1..414
                /organism="Sterkiella histriomuscorum"
                /mol_type="genomic DNA"
                /db_xref="taxon:94289"
                /clone="UUGC10005D19"
                /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                /clone_lib="Oxytricha plasmid UUGC10 library"
                /note="vector: PWD42nv; Purified macronuclear chromosomal
                DNA from Oxytricha trifallax was blunt end-repaired with
                T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
                oligonucleotides were ligated to the blunt ends in high
                molar excess. Vector DNA was prepared from a derivative of
                PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
                derivative of plasmid R1. The vector was ligated with
                adaptors complementary to the insert adaptors and
                purified. The sheared, adapted mouse DNA was annealed to
                adapted vector DNA, and transformed into
                chemically-competent E. Coli XL10-Gold (Stratagene) cells
                and selected for ampicillin resistance."
ORIGIN
Query Match      100.0%; Score 101; DB 9; Length 414;
Best Local Similarity 100.0%; Pred. No. 8.3e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAACAATAG 60
Db 414 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAACAATAG 355
Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
Db 354 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 314

```

```

RESULT 7
BJ684174/c
LOCUS
DEFINITION BJ684174 HCEST library Haplochromis chilotes cDNA clone no90c12,
            mRNA sequence.
ACCESSION BJ684174
VERSION BJ684174.1 GI:46527295
KEYWORDS EST
SOURCE Haplochromis chilotes
        ORGANISM Haplochromis chilotes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Acanthopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes;
            Labroidae; Cichlidae; Haplochromis.
REFERENCE 1 (bases 1 to 417)
AUTHORS Watanabe,M., Kobayashi,N., Shin-I,T., Kohara,Y. and Okada,N.
TITLE Orf sequences of cichlid in Lake Victoria are essentially same
JOURNAL Unpublished (2004)
COMMENT Contact: Tadasu Shin-i
        Center For Genetic Resource Information
        National Institute of Genetics
        1111 Yata, Mishima, Shizuoka 411-8540, Japan
        Tel: 81-559-81-6856
        Fax: 81-559-81-6855
        Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source           1..417
                     /organism="Haplochromis chilotes"
                     /mol_type="mRNA"
                     /db_xref="taxon:257977"
                     /clone="no90c12"
                     /tissue_type="jaw"
                     /dev_stage="varied"
                     /clone_lib="HCEST library"

ORIGIN
Query Match      100.0%; Score 101; DB 4; Length 417;
Best Local Similarity 100.0%; Pred. No. 8.3e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTAGTATTAGAAAAATAACAAATAG 60
    |||||
Db 129 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTAGTATTAGAAAAATAACAAATAG 70
    |||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101
    |||||
Db 69 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 29
    |||||

RESULT 8
CC819923/c
LOCUS
DEFINITION CC819923 Oxytricha plasmid UUGC10 library Sterkiella
            histriomuscorum genomic clone UUGC10006J13 R, genomic survey
            sequence.
ACCESSION CC819923
VERSION CC819923.1 GI:32900671
KEYWORDS GSS.
SOURCE Sterkiella histriomuscorum (Oxytricha trifallax)
        ORGANISM Sterkiella histriomuscorum
            Eukaryota; Alveolata; Ciliophora; Spirotriches; Stichotrichia;
            Stichotrichida; Oxytrichidae; Sterkiella.
REFERENCE 1 (bases 1 to 491)
AUTHORS Dunn,D., Doak,T., Herrick,G. and Weiss,R.
TITLE Paired end reads from plasmid inserts of Oxytricha trifallax
JOURNAL Unpublished (2003)
COMMENT Contact: Robert B. Weiss
        University of Utah Genome Center
        University of Utah
        Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
        84112, USA
        Tel: 801 585 5606

```

```

Fax: 801 585 7177
Email: dunn@genetics.utah.edu
Plate: 0006 row: J column: 13
Seq primer: CACACAGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 491.
FEATURES             Location/Qualifiers
     source           1..491
                     /organism="Sterkiella histriomuscorum"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:94289"
                     /clone="UUGC10006J13"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /notes="Oxytricha plasmid UUGC10 library"
                     /clone_lib="Oxytricha plasmid UUGC10 library"
                     /notes="Vector: PWD42nv; Purified macronuclear chromosomal
                     DNA from Oxytricha trifallax was blunt end-repaired with
                     T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
                     oligonucleotides were ligated to the blunt ends in high
                     molar excess. Vector DNA was prepared from a derivative of
                     pWD42 (GI4732114|9b|AF129072.1), a copy-number inducible
                     derivative of plasmid R1. The vector was ligated with
                     adaptors complementary to the insert adaptors and
                     purified. The sheared, adapted mouse DNA was annealed to
                     adapted vector DNA, and transformed into
                     chemically-competent E. Coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."

ORIGIN
Query Match      100.0%; Score 101; DB 9; Length 491;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTAGTATTAGAAAAATAACAAATAG 60
    |||||
Db 412 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTAGTATTAGAAAAATAACAAATAG 353
    |||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101
    |||||
Db 352 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 312
    |||||

RESULT 9
BI805285
LOCUS
DEFINITION BI805285 Stem library from Oryza sativa (3-5 leaf stage) Oryza
            sativa cDNA clone S035A01, mRNA sequence.
ACCESSION BI805285
VERSION BI805285.1 GI:15852489
KEYWORDS EST.
SOURCE Oryza sativa
        ORGANISM Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 495)
AUTHORS Dong,H.T., Li,D.B., Zhuang,X.F., Dai,C.G., Sun,L.X., Pei,Y.X.,
            Wu,H.F., Jiang,Y.X., Yu,F.C., Gao,Q.K. and Lou,Y.C.
TITLE A Gene Expression Screen in Oryza sativa
JOURNAL Unpublished (2001)
COMMENT Contact: Haitao Dong, Dehao Li
        Bioinformatics and Gene Network Research Group
        Zhejiang University
        Kaixuan Road 268#, Hangzhou, Zhejiang, P.R.China
        Tel: 0086-571-86892051
        Fax: 0086-571-86961525
        Email: webmaster@estarray.org, URL: http://www.estarray.org
        Seq primer: M13 forward primer.
FEATURES             Location/Qualifiers
     source           1..495
                     /organism="Oryza sativa"
                     /mol_type="mRNA"
                     /db_xref="taxon:4530"
                     /clone="S035A01"

```

```
/tissue_type="Stem"
/dev_stage="3-5 leaf stage"
/clone_lib="Stem library from Oryza sativa (3-5 leaf
stage)"
/note="Vector: pSport2"

ORIGIN

Query Match          100.0%; Score 101; DB 4; Length 495;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAATAAACAATAG 60
    |||||||
Db 62 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAATAAACAATAG 121

Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
    |||||||
Db 122 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 162

RESULT 10
CC818374          495 bp      DNA      linear      GSS 17-JUL-2003
LOCUS             100004807R Oxytricha plasmid UUGC100004807 R, genomic survey
DEFINITION        histriomuscorum genomic clone UUGC100004807 R, genomic survey
sequence.
ACCESSION         CC818374
VERSION           CC818374.1 GI:32897661
KEYWORDS          GSS.
SOURCE            Sterkiella histriomuscorum (Oxytricha trifallax)
ORGANISM          Sterkiella histriomuscorum
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
Stichotrichida; Oxytrichidae; Sterkiella.
REFERENCE         1 (bases 1 to 495)
AUTHORS           Dunn, D., Doak, T., Herrick, G. and Weiss, R.
TITLE             Paired end reads from plasmid inserts of Oxytricha trifallax
macronuclear chromosomes
JOURNAL           Unpublished (2003)
COMMENT           Contact: Robert B. Weiss
                  University of Utah Genome Center
                  Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
                  84112, USA
                  Tel: 801 585 5606
                  Fax: 801 585 7177
                  Email: ddunn@genetics.utah.edu
                  Plate: 0004 row: B column: 07
                  Seq primer: CACACGAGAAACAGCTATGACC
                  Class: plasmid ends
                  High quality sequence stop: 495.
FEATURES          Location/Qualifiers
source            1..495
                  /organism="Sterkiella histriomuscorum"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:94289"
                  /clone="UUGC100004807"
                  /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
                  /clone_lib="Oxytricha plasmid UUGC10 library"
                  /note="Vector: PWD42nv; Purified macronuclear chromosomal
                  DNA from Oxytricha trifallax was blunt end-repaired with
                  T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
                  oligonucleotides were ligated to the blunt ends in high
                  molar excess. Vector DNA was prepared from a derivative of
                  PWD42 [G14732114|gb|AF129072.1], a copy-number inducible
                  derivative of plasmid R1. The vector was ligated with
                  adaptors complementary to the insert adaptors and
                  purified. The sheared, adapted mouse DNA was annealed to
                  chemically-competent E. Coli XL10-Gold (Stratagene) cells
                  and selected for ampicillin resistance."

ORIGIN

Query Match          100.0%; Score 101; DB 9; Length 495;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAATAAACAATAG 60
    |||||||
Db 391 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAATAAACAATAG 332

Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
    |||||||
```

```
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAATAAACAATAG 60
    |||||||
Db 392 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAATAAACAATAG 333

Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
    |||||||
Db 332 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 292

RESULT 11
CC818523          496 bp      DNA      linear      GSS 17-JUL-2003
LOCUS             100004L13R Oxytricha plasmid UUGC10 library Sterkiella
DEFINITION        histriomuscorum genomic clone UUGC100004L13 R, genomic survey
sequence.
ACCESSION         CC818523
VERSION           CC818523.1 GI:32897943
KEYWORDS          GSS.
SOURCE            Sterkiella histriomuscorum (Oxytricha trifallax)
ORGANISM          Sterkiella histriomuscorum
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
Stichotrichida; Oxytrichidae; Sterkiella.
REFERENCE         1 (bases 1 to 496)
AUTHORS           Dunn, D., Doak, T., Herrick, G. and Weiss, R.
TITLE             Paired end reads from plasmid inserts of Oxytricha trifallax
macronuclear chromosomes
JOURNAL           Unpublished (2003)
COMMENT           Contact: Robert B. Weiss
                  University of Utah Genome Center
                  Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
                  84112, USA
                  Tel: 801 585 5606
                  Fax: 801 585 7177
                  Email: ddunn@genetics.utah.edu
                  Plate: 0004 row: L column: 13
                  Seq primer: CACACGAGAAACAGCTATGACC
                  Class: plasmid ends
                  High quality sequence stop: 496.
FEATURES          Location/Qualifiers
source            1..496
                  /organism="Sterkiella histriomuscorum"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:94289"
                  /clone="UUGC100004L13"
                  /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
                  /clone_lib="Oxytricha plasmid UUGC10 library"
                  /note="Vector: PWD42nv; Purified macronuclear chromosomal
                  DNA from Oxytricha trifallax was blunt end-repaired with
                  T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
                  oligonucleotides were ligated to the blunt ends in high
                  molar excess. Vector DNA was prepared from a derivative of
                  PWD42 [G14732114|gb|AF129072.1], a copy-number inducible
                  derivative of plasmid R1. The vector was ligated with
                  adaptors complementary to the insert adaptors and
                  purified. The sheared, adapted mouse DNA was annealed to
                  chemically-competent E. Coli XL10-Gold (Stratagene) cells
                  and selected for ampicillin resistance."

ORIGIN

Query Match          100.0%; Score 101; DB 9; Length 496;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAATAAACAATAG 60
    |||||||
Db 391 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAATAAACAATAG 332

Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
    |||||||
```


REFERENCE 1 (bases 1 to 518)
 AUTHORS Dunn,D., Doak,T., Herrick,G. and Weiss,R.
 TITLE Paired end reads from plasmid inserts of *Oxytricha trifallax*
 JOURNAL macronuclear chromosomes
 COMMENT Unpublished (2003)
 CONTACT: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: dunn@genetics.utah.edu
 Plate: 0002 row: D column: 21
 Seq primer: CACACAGGAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 518.

FEATURES
 source
 1..518
 Location/Qualifiers
 /organism="Sterkiella histriomuscorum"
 /mol_type="genomic DNA"
 /db_xref="taxon:94289"
 /clone="UUGC100002D21"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /clone_lib="Oxytricha plasmid UUGC10 library"
 /note="Vector: PWD42nv; Purified macronuclear chromosomal DNA from *Oxytricha trifallax* was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. Vector DNA was prepared from a derivative of pMD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. Coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN
 Query Match 100.0%; Score 101; DB 9; Length 518;
 Best Local Similarity 100.0%; Pred. No. 8.4e-19;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 AGGGTTATTGCTCATGCGGATACATATTTGAATGTATTAGAAAAATAACAAATAG 60
 Db 410 AGGGTTATTGCTCATGCGGATACATATTTGAATGTATTAGAAAAATAACAAATAG 351
 Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
 Db 350 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 310

RESULT 15
 CC817162/c
 LOCUS CC817162.1
 DEFINITION 519 bp DNA linear GSS 17-JUL-2003
 histriomuscorum genomic clone UUGC100002J19 R, genomic survey sequence.
 CC817162
 CC817162.1 GI:32896449
 GSS.
 SOURCE Sterkiella histriomuscorum (*Oxytricha trifallax*)
 ORGANISM
 Sterkiella histriomuscorum
 Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
 Stichotrichida; Oxytrichidae; Sterkiella.
 1 (bases 1 to 519)
 Dunn,D., Doak,T., Herrick,G. and Weiss,R.
 Paired end reads from plasmid inserts of *Oxytricha trifallax*
 macronuclear chromosomes
 UNPUBLISHED (2003)
 CONTACT: Robert B. Weiss
 University of Utah Genome Center

University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: dunn@genetics.utah.edu
 Plate: 0002 row: J column: 19
 Seq primer: CACACAGGAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 519.

FEATURES
 source
 1..519
 Location/Qualifiers
 /organism="Sterkiella histriomuscorum"
 /mol_type="genomic DNA"
 /db_xref="taxon:94289"
 /clone="UUGC100002J19"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /clone_lib="Oxytricha plasmid UUGC10 library"
 /note="Vector: PWD42nv; Purified macronuclear chromosomal DNA from *Oxytricha trifallax* was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. Vector DNA was prepared from a derivative of pMD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. Coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN
 Query Match 100.0%; Score 101; DB 9; Length 519;
 Best Local Similarity 100.0%; Pred. No. 8.4e-19;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 AGGGTTATTGCTCATGCGGATACATATTTGAATGTATTAGAAAAATAACAAATAG 60
 Db 416 AGGGTTATTGCTCATGCGGATACATATTTGAATGTATTAGAAAAATAACAAATAG 357
 Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
 Db 356 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 316

Search completed: July 14, 2005, 23:23:06
 Job time : 962.667 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:39:07 ; Search time 749.127 Seconds
(without alignments)
6468.225 Million cell updates/sec

Title: US-09-482-682-64_COPY_1_100
Perfect score: 100
Sequence: 1 ctgctccctgtgtgtgtt.....caattgatgaagaattctgc 100

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.on.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100	100.0	562	6	AX643583 Sequence
2	100	100.0	633	14	ALRPROLTB
3	100	100.0	648	6	AX175190 Sequence
4	100	100.0	648	6	AX175195 Sequence
5	100	100.0	1070	6	A85308 Sequence
6	100	100.0	1070	6	BD107647
7	100	100.0	2245	6	AX643582 Sequence
8	100	100.0	2426	6	AX044426 Sequence
9	100	100.0	2427	6	AX044425 Sequence
10	100	100.0	3557	12	SYNRSV3MV
11	100	100.0	3840	12	EVEI32038
12	100	100.0	3853	6	AR098190 Sequence
13	100	100.0	3853	6	AR207832 Sequence
14	100	100.0	3853	6	BD009729 Tissue sp
15	100	100.0	3925	6	A60213 Sequence
16	100	100.0	3925	6	AR122289 Sequence
17	100	100.0	3986	12	PCDNA32EO
18	100	100.0	4026	6	AR098191 Sequence
19	100	100.0	4026	6	AR207833 Sequence

20	100	100.0	4026	6	BD009730
c	21	100	100.0	6	AR071324
	22	100	100.0	6	AR098192 Sequence
	23	100	100.0	6	AR207834 Sequence
	24	100	100.0	6	BD009731 Tissue sp
	25	100	100.0	6	A38214
	26	100	100.0	6	AX286570 Sequence
	27	100	100.0	6	AX743954 Sequence
	28	100	100.0	6	AR062871 Sequence
	29	100	100.0	6	AX060344 Sequence
	30	100	100.0	12	SYNRSV5GPT
	31	100	100.0	6	AX133940 Sequence
	32	100	100.0	6	AR071323 Sequence
	33	100	100.0	6	BD238492 Expressio
	34	100	100.0	6	AX234391 Sequence
	35	100	100.0	6	A91754
	36	100	100.0	6	BD085110 Vertebrat
	37	100	100.0	12	SYNRSV5NEO
	38	100	100.0	6	AX951626 Sequence
	39	100	100.0	12	CVU89673
	40	100	100.0	6	BD234590 Screening
	41	100	100.0	6	AX026821 Sequence
	42	100	100.0	6	BD195386 Compositi
	43	100	100.0	6	AX319694 Sequence
	44	100	100.0	12	SYNFCRC
	45	100	100.0	6	A44171 Sequence 1

ALIGNMENTS

RESULT 1
AX643583
LOCUS AX643583 562 bp DNA linear PAT 24-FEB-2003
DEFINITION Sequence 2 from Patent WO02099100.
ACCESSION AX643583
VERSION AX643583.1 GI:28551383
KEYWORDS
SOURCE Mus 'sp'.
ORGANISM Mus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Al-Rubeai, M. and Shuttleworth, J.
TITLE Method of production of a protein in cells which inducibly express the cell cycle inhibitor protein, p21
JOURNAL Patent: WO 02099100-A 2 12-DEC-2002;
Lonza Biologics plc (GB)
FEATURES
source Location/Qualifiers
1..562
/organism="Mus sp."
/mol_type="unassigned DNA"
/db_xref="taxon:10095"
/note="Rous Sarcoma Virus LTR promoter"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 562;
Best Local Similarity 100.0%; Pred. No. 2.6e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTGCTCCCTGCTGTGTGTGGAGTGCCTGAGTAGTGCAGCAAAATTAAGCTACA 60
Db 46 CTGCTCCCTGCTGTGTGTGGAGTGCCTGAGTAGTGCAGCAAAATTAAGCTACA 105
Qy 61 ACAAGCAAGCTTGCACCGCAATTCATGAAGAAATCTGC 100
Db 106 ACAAGCAAGCTTGCACCGCAATTCATGAAGAAATCTGC 145

RESULT 2
ALRPROLTB
LOCUS ALRPROLTB 633 bp ss-RNA linear VRL 28-APR-1993
DEFINITION Rous sarcoma virus (Schmidt-Ruppin), proviral, 3' LTR on 21S mRNA.

J02025 J02022
 J02025.1 GI:210255
 C-myc proto-oncogene; long terminal repeat (LTR); src oncogene.
 SOURCE Rous sarcoma virus
 ORGANISM Rous sarcoma virus
 Viruses; Retrovirdae; Retroviridae; Alpharetrovirus.
 REFERENCE 1 (sites)
 Yamanoto,T., de Crombrughe,B. and Pastan,I.
 Identification of a functional promoter in the long terminal repeat
 of Rous sarcoma virus
 JOURNAL Cell 22 (3), 787-797 (1980)
 MEDLINE 81112147
 PUBMED 6257399
 REFERENCE 2 (bases 1 to 633)
 Yamanoto,T., Tyagi,J.S., Pagan,J.B., Jay,G., deCrombrughe,B. and
 Pastan,I.
 Molecular mechanism for the capture and excision of the
 transforming gene of avian sarcoma virus as suggested by analysis
 of recombinant clones
 J. Virol. 35 (2), 436-443 (1980)
 JOURNAL 81072438
 MEDLINE 6255184
 PUBMED 3 (bases 319 to 633)
 REFERENCE Yamanoto,T., Jay,G. and Pastan,I.
 Unusual features in the nucleotide sequence of a cDNA clone derived
 from the common region of avian sarcoma virus messenger RNA
 Proc. Natl. Acad. Sci. U.S.A. 77 (1), 176-180 (1980)
 JOURNAL 80145590
 MEDLINE 6244542
 PUBMED Original source text: Rous sarcoma virus (Schmidt-Ruppin strain,
 subgroup D) provirus, cDNA to 21S mRNA from infected chicken
 embryonic fibroblasts, clone pSR1.
 COMMENT [1] sites; mRNA start.
 Original figure in [2] included 24 'g's on 5' end and 16 'c's on 3'
 end that were cDNA synthesis artifacts.
 [2] also sequenced a defective clone, pSR2, with the src gene
 deleted (see separate entry).
 [1] demonstrated the mRNA transcription initiation site shown in
 the Sites table using pSR1 as a template. However, this is the 3',
 LTR, and the functional mRNA start site would be assumed to be on
 the 5' LTR at the homologous site.
 FEATURES Location/Qualifiers
 source 1..633
 /organism="Rous sarcoma virus"
 /mol_type="genomic RNA"
 /db_xref="taxon:11886"
 misc_RNA <1..517
 /note="viroion genomic RNA"
 LTR 211..5633
 /note="3' LTR"
 mRNA 517..5633
 /note="in vitro mRNA [1]; see comment"
 repeat_region 517..536
 /note="terminally redundant repeat"
 ORIGIN 20 bp upstream of pSR1 site.
 Query Match 100.0%; Score 100; DB 14; Length 633;
 Best Local Similarity 100.0%; Pred. No. 2.6e-25;
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 CTGCTCCCTGCTGTGTGGAGTGCCTGCTAGTAGTCGCGAGCAAAATTTAAGCTACA 60
 Db 28 CTGCTCCCTGCTGTGTGTGGAGTGCCTGCTAGTAGTCGCGAGCAAAATTTAAGCTACA 87
 Qy 61 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 100
 Db 88 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 127
 RESULT 3
 AX175190 648 bp DNA linear PAT 03-JUL-2001
 LOCUS Sequence 1 from Patent WO0142444.
 DEFINITION

AX175190
 AX175190.1 GI:14598581
 synthetic construct
 SOURCE synthetic construct
 ORGANISM synthetic construct
 other sequences; artificial sequences.
 REFERENCE 1
 Rivera,V., Zoltick,P. and Wilson,J.M.
 Methods for expression of genes in primates
 TITLE Patent: WO 0142444-A 1 14-JUN-2001;
 JOURNAL ARIAD GENE THERAPEUTICS, INC. (US) ; THE UNIVERSITY OF PENNSYLVANIA
 (US)
 FEATURES Location/Qualifiers
 source 1..648
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"
 /note="vector/RSV promoter/vector"
 ORIGIN
 Query Match 100.0%; Score 100; DB 6; Length 648;
 Best Local Similarity 100.0%; Pred. No. 2.6e-25;
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 CTGCTCCCTGCTGTGTGTGGAGTGCCTGCTAGTAGTCGCGAGCAAAATTTAAGCTACA 60
 Db 90 CTGCTCCCTGCTGTGTGTGGAGTGCCTGCTAGTAGTCGCGAGCAAAATTTAAGCTACA 149
 Qy 61 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 100
 Db 150 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 189
 RESULT 4
 AX175195 648 bp DNA linear PAT 03-JUL-2001
 LOCUS Sequence 6 from Patent WO0142444.
 DEFINITION
 AX175195
 AX175195
 AX175195.1 GI:14598586
 synthetic construct
 SOURCE synthetic construct
 ORGANISM other sequences; artificial sequences.
 REFERENCE 1
 Rivera,V., Zoltick,P. and Wilson,J.M.
 Methods for expression of genes in primates
 TITLE Patent: WO 0142444-A 6 14-JUN-2001;
 JOURNAL ARIAD GENE THERAPEUTICS, INC. (US) ; THE UNIVERSITY OF PENNSYLVANIA
 (US)
 FEATURES Location/Qualifiers
 source 1..648
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"
 /note="MluI/RSV promoter/BglI"
 ORIGIN
 Query Match 100.0%; Score 100; DB 6; Length 648;
 Best Local Similarity 100.0%; Pred. No. 2.6e-25;
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 CTGCTCCCTGCTGTGTGTGGAGTGCCTGCTAGTAGTCGCGAGCAAAATTTAAGCTACA 60
 Db 90 CTGCTCCCTGCTGTGTGTGGAGTGCCTGCTAGTAGTCGCGAGCAAAATTTAAGCTACA 149
 Qy 61 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 100
 Db 150 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 189
 RESULT 5
 A85308 1070 bp DNA linear PAT 21-JAN-2000
 LOCUS

```
DEFINITION Sequence 6 from Patent WO9840493.
ACCESSION A85308
VERSION A85308.1 GI:6733916
KEYWORDS unidentified
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1070)
AUTHORS Rigby,M.A. and Jarrett,J.O.
TITLE FIV VACCINE
JOURNAL Patent: WO 9840493-A 6 17-SEP-1998;
RIGBY MARK ALAN (GB); JARRETT JAMES OSWALD (GB)
FEATURES
    source
        Location/Qualifiers
            1..1070
                /organism="unidentified"
                /mol_type="unassigned DNA"
                /db_xref="taxon:32644"
ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 1070;
Best Local Similarity 100.0%; Pred. No. 2.8e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTGCTCCCTGCTGTGTGGAGTGCCTGAGTAGTGGCGGAGCAAAATTTAAGCTACA 60
Db 77 CTGCTCCCTGCTGTGTGTGGAGTGCCTGAGTAGTGGCGGAGCAAAATTTAAGCTACA 136
Qy 61 ACAAGGCAAGGCTTGACCGACAAATTCATGAAGAATCTGC 100
Db 137 ACAAGGCAAGGCTTGACCGACAAATTCATGAAGAATCTGC 176
RESULT 7
AX0443582 LOCUS AX643582 2245 bp DNA linear PAT 24-FEB-2003
DEFINITION Sequence 1 from Patent WO02099100.
ACCESSION AX643582
VERSION AX643582.1 GI:28551382
KEYWORDS
SOURCE Mus sp.
ORGANISM Mus sp.
REFERENCE 1
AUTHORS Al-Rubeai,M. and Shuttleworth,J.
TITLE Method of production of a protein in cells which inducibly express
the cell cycle inhibitor protein, p21
JOURNAL Patent: WO 02099100-A 1 12-DEC-2002;
Lonza Biologics plc (GB)
FEATURES
    source
        Location/Qualifiers
            1..2245
                /organism="Mus sp."
                /mol_type="unassigned DNA"
                /db_xref="taxon:10095"
                /note="RSV-LTR promoter + intron + p21 cds + Tkpoly(A)
                LacSwitch II expression construct"
ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 2245;
Best Local Similarity 100.0%; Pred. No. 3e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTGCTCCCTGCTGTGTGGAGTGCCTGAGTAGTGGCGGAGCAAAATTTAAGCTACA 60
Db 46 CTGCTCCCTGCTGTGTGTGGAGTGCCTGAGTAGTGGCGGAGCAAAATTTAAGCTACA 105
Qy 61 ACAAGGCAAGGCTTGACCGACAAATTCATGAAGAATCTGC 100
Db 106 ACAAGGCAAGGCTTGACCGACAAATTCATGAAGAATCTGC 145
RESULT 8
AX044426 LOCUS AX044426 2426 bp DNA linear PAT 24-NOV-2000
DEFINITION Sequence 18 from Patent WO0066752.
ACCESSION AX044426
VERSION AX044426.1 GI:11343299
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Castro,M.G., Emery,S.C. and Lowenstein,P.R.
TITLE Chemical compounds
JOURNAL Patent: WO 0066752-A 18 09-NOV-2000;
Astrazeneca AB (SE); THE VICTORIA UNIVERSITY OF MANCHESTER (GB)
FEATURES
    source
        Location/Qualifiers
            1..2426
                /organism="synthetic construct"
                /mol_type="unassigned DNA"
                /db_xref="taxon:32630"
                /note="CPG2 with last exon of Thy-1 fused at 3' end"
ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 1070;
Best Local Similarity 100.0%; Pred. No. 2.8e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTGCTCCCTGCTGTGTGGAGTGCCTGAGTAGTGGCGGAGCAAAATTTAAGCTACA 60
Db 77 CTGCTCCCTGCTGTGTGTGGAGTGCCTGAGTAGTGGCGGAGCAAAATTTAAGCTACA 136
Qy 61 ACAAGGCAAGGCTTGACCGACAAATTCATGAAGAATCTGC 100
Db 137 ACAAGGCAAGGCTTGACCGACAAATTCATGAAGAATCTGC 176
RESULT 6
BD107647 LOCUS BD107647 1070 bp DNA linear PAT 18-SEP-2002
DEFINITION FIV vaccine.
ACCESSION BD107647
VERSION BD107647.1 GI:23202465
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 1070)
AUTHORS Neil,J.C., Rigby,M.A. and Jarrett,J.O.
TITLE FIV vaccine
JOURNAL Patent: JP 2002501369-A 6 15-JAN-2002;
THE UNIVERSITY COURT OF THE UNIVERSITY OF GLASGOW
COMMENT OS Artificial Sequence
PN JP 2002501369-A/6
PD 15-JAN-2002
PF 10-MAR-1998 JP 1998539351
PR 11-MAR-1997 GB 9704977.9
PI JAMES CHARLES NEIL,MARK ALAN RIGBY,JAMES OSWALD JARRETT PC
C12N15/49,A61K31/70,A61K48/00
CC CMV PROMOTER FROM pCDNA3 (a Bgl II - Kpn
I restriction fragment)
CC SST I - SST I FRAGMENT IN PLASMID CMV DEL. RT CC FIV GENOME
FROM THE t-RNA PRIMER BINDING
SITE TO THE VIRAL SAT
CC IS
FH Key Location/Qualifiers
FT source 1..1070
    /organism="Artificial Sequence".
FEATURES
    source
        Location/Qualifiers
            1..1070
                /organism="synthetic construct"
                /mol_type="genomic DNA"
                /db_xref="taxon:32630"
ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 1070;
Best Local Similarity 100.0%; Pred. No. 2.8e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTGCTCCCTGCTGTGTGGAGTGCCTGAGTAGTGGCGGAGCAAAATTTAAGCTACA 60
Db 77 CTGCTCCCTGCTGTGTGTGGAGTGCCTGAGTAGTGGCGGAGCAAAATTTAAGCTACA 136
Qy 61 ACAAGGCAAGGCTTGACCGACAAATTCATGAAGAATCTGC 100
Db 137 ACAAGGCAAGGCTTGACCGACAAATTCATGAAGAATCTGC 176
```

Query Match 100.0%; Score 100; DB 6; Length 2426;
Best Local Similarity 100.0%; Pred. No. 3.1e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTTGTGTTGGAGTCTGCTGAGTAGTGCAGCAAAATTTAAGCTACA 60
|||||
Db 69 CTGCTCCCTGCTTGTGTTGGAGTCTGCTGAGTAGTGCAGCAAAATTTAAGCTACA 128
|||||

Qy 61 ACAAGGCAAGCTTGACCGACAATTGCATGAAGAATCTGC 100
|||||
Db 129 ACAAGGCAAGCTTGACCGACAATTGCATGAAGAATCTGC 168
|||||

RESULT 9
LOCUS AX044425 2427 bp DNA linear PAT 24-NOV-2000
DEFINITION Sequence 17 from Patent WO0066752.
ACCESSION AX044425
VERSION AX044425.1 GI:11343298
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.

REFERENCE 1
AUTHORS Castro, M.G., Emery, S.C. and Lowenstein, P.R.
TITLE Chemical compounds
JOURNAL Patent: WO 0066752-A 17 09-NOV-2000;
Astrazeneca AB (SE); THE VICTORIA UNIVERSITY OF MANCHESTER (GB)

FEATURES
source
1. 2427
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="CPG2 mutant with last exon of Thy-1 fused at 3'
end"

ORIGIN

Query Match 100.0%; Score 100; DB 6; Length 2427;
Best Local Similarity 100.0%; Pred. No. 3.1e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTTGTGTTGGAGTCTGCTGAGTAGTGCAGCAAAATTTAAGCTACA 60
|||||
Db 70 CTGCTCCCTGCTTGTGTTGGAGTCTGCTGAGTAGTGCAGCAAAATTTAAGCTACA 129
|||||

Qy 61 ACAAGGCAAGCTTGACCGACAATTGCATGAAGAATCTGC 100
|||||
Db 130 ACAAGGCAAGCTTGACCGACAATTGCATGAAGAATCTGC 169
|||||

RESULT 10
LOCUS SYNRSV3MV 3557 bp DNA circular SYN 27-APR-1993
DEFINITION Cloning vector RSV3.
ACCESSION M83240
VERSION M83240.1 GI:209303
KEYWORDS cDNA expression vector.
SOURCE unidentified cloning vector
ORGANISM other sequences; artificial sequences; vectors.

REFERENCE 1
AUTHORS Messing, J.
TITLE New M13 vectors for cloning
JOURNAL Meth. Enzymol. 101, 20-78 (1983)
MEDLINE 83296918
PUBMED 6310323

REFERENCE 2
AUTHORS Gorman, C., Padmanabhan, R. and Howard, B.H.
TITLE High efficiency DNA-mediated transformation of primate cells
JOURNAL Science 221 (4610), 551-553 (1983)
MEDLINE 83249156
PUBMED 6306768

REFERENCE 3
AUTHORS (bases 1 to 3557)

AUTHORS Jacobson, S., Sekaly, R.P., Jacobson, C.L., McFarland, H.F. and Long, B.O.
TITLE HUA class II-restricted presentation of cytoplasmic measles virus antigens to cytotoxic T cells
JOURNAL J. Virol. 63 (4), 1756-1762 (1989)
MEDLINE 89178863
PUBMED 2784508
COMMENT Original source text: Cloning vector DNA.
FEATURES
source
1. 3557
Location/Qualifiers
/organism="unidentified cloning vector"
/mol_type="genomic DNA"
/db_xref="taxon:45196"
misc_feature
1. 29
/function="polylinker"
/evidence="experimental"
misc_feature
912..3029
/function="ampicillin-resistance, replication origin"
/evidence="experimental"
enhancer
3030..3557
/standard_name="5'LTR of Rous Sarcoma Virus"
/citation=[2]
/evidence="experimental"

ORIGIN

Query Match 100.0%; Score 100; DB 12; Length 3557;
Best Local Similarity 100.0%; Pred. No. 3.2e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTTGTGTTGGAGTCTGCTGAGTAGTGCAGCAAAATTTAAGCTACA 60
|||||
Db 3030 CTGCTCCCTGCTTGTGTTGGAGTCTGCTGAGTAGTGCAGCAAAATTTAAGCTACA 3089
|||||

Qy 61 ACAAGGCAAGCTTGACCGACAATTGCATGAAGAATCTGC 100
|||||
Db 3090 ACAAGGCAAGCTTGACCGACAATTGCATGAAGAATCTGC 3129
|||||

RESULT 11
LOCUS EVE132038 3840 bp RNA circular SYN 28-JUL-1999
DEFINITION Expression vector pCDPT.
ACCESSION AJ132038
VERSION AJ132038.1 GI:5640088
KEYWORDS AMP gene; beta lactamase; Cole1 origin of replication; multiple cloning site; SP6 promoter; SV40 origin of replication; T7 promoter; xanthine-guanine phosphoribosyl transferase; xanthine-guanine phosphoribosyl transferase gene.

SOURCE Expression vector pCDPT
ORGANISM Expression vector pCDPT
other sequences; artificial sequences; vectors.

REFERENCE 1
AUTHORS Zeng, B.J.
TITLE Mammalian Expression Vector for with fuse Xanthine-guanine phosphoribosyl transferase Tag
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3840)
AUTHORS Zeng, B.J.
TITLE Direct Submission
JOURNAL Submitted (27-FEB-1999) Zeng B.J., Gene Engineering Center, Institute of Microbiology, Zhongguancun, Beijing, Beijing 100080, CHINA

FEATURES
source
1. 3840
Location/Qualifiers
/organism="Expression vector pCDPT"
/mol_type="other RNA"
/db_xref="taxon:90749"
promoter
209..863
/note="CMV"
promoter
864..882
/note="T7"
misc_feature
882..984
/note="Multiple cloning site"

```

CDS
HindIII, BamHI, BstXI, EcoRI, NotI, XhoI
929..1387
/codon_start=1
/product="Xanthine-guanine phosphoribosyl transferase"
/protein_id="CAB51567.1"
/db_xref="GI:5640089"
/translation="MSKSYIVTWMLQIHARKKLASRLMPSEOWKGIITAVSRGGLVPGA
LLARELGHVHDVTVICISYDHNORELKLKRAEGDGEFIVIDDLVDVTGTAVALRE
MYPKAFVTTTIPAKPAGRLVDDYVVDIPQDTWIEQPDGMGVFVPPISGR"
1649..1863
/feature="BGN"
2450..2775
/feature="SP6"
2644..2729
/feature="SV40"
complement(2844..3704)
/feature="amp"
complement(2844..3704)
/feature="amp"
/codon_start=1
/product="beta-lactamase"
/protein_id="CAB51568.1"
/db_xref="GI:5640090"
/translation="MSIQHFRVALIPFAHPECLPVFAHPETLVKVKDAEDQLGARVGY
IELDLSGKILDSRPERFPMWTFKLLGAVLSRIDAGQEQLRRIRHYSQNDLVE
YSPVTEKLTIDGMTRELCSAAITMSDNTAANLLLTITGGPKELTAFHNNGDHVTSL
DRWPELNEATPDERDTTMEVAMATTLKLLTGLLTLLASRQQLIDWMEADKVGPL
LRSALPAGWFTADKSGAGERSGRIIAALGPDGPKSRIVVIYTTGSGQTWDERNRQIA
EIGASLIKHW"
3632..3840
/feature="ColEI"

rep_origin
Query Match 100.0%; Score 100; DB 12; Length 3840;
Best Local Similarity 100.0%; Pred. No. 3.3e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTTGTGTGTGGAGTCTGCTGAGTAGTGCAGCAAAATTAAGCTACA 60
Db 6 CTGCTCCCTGCTTGTGTGTGGAGTCTGCTGAGTAGTGCAGCAAAATTAAGCTACA 65

Qy 61 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 100
Db 66 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 105

RESULT 12
AR098190 LOCUS 3853 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 5 from patent US 6074850.
ACCESSION AR098190
VERSION AR098190.1 GI:12807447
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3853)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion polypeptides
JOURNAL Patent: US 6074850-A 5 13-JUN-2000;
FEATURES Location/Qualifiers
source 1..3853
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 3853;
Best Local Similarity 100.0%; Pred. No. 3.3e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTTGTGTGTGGAGTCTGCTGAGTAGTGCAGCAAAATTAAGCTACA 60
Db 81 CTGCTCCCTGCTTGTGTGTGGAGTCTGCTGAGTAGTGCAGCAAAATTAAGCTACA 140

Qy 61 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 100
Db 66 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 105

RESULT 13
AR207832 LOCUS 3853 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 5 from patent US 6379927.
ACCESSION AR207832
VERSION AR207832.1 GI:21507688
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3853)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion proteins
JOURNAL Patent: US 6379927-A 5 30-APR-2002;
FEATURES Location/Qualifiers
source 1..3853
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 3853;
Best Local Similarity 100.0%; Pred. No. 3.3e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTTGTGTGTGGAGTCTGCTGAGTAGTGCAGCAAAATTAAGCTACA 60
Db 81 CTGCTCCCTGCTTGTGTGTGGAGTCTGCTGAGTAGTGCAGCAAAATTAAGCTACA 140

Qy 61 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 100
Db 141 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 180

RESULT 14
BD009729 LOCUS 3853 bp DNA linear PAT 31-JAN-2002
DEFINITION Tissue specific expression of retinoblastoma protein.
ACCESSION BD009729
VERSION BD009729.1 GI:18638102
KEYWORDS JP 2001503638-A/3.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 3853)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Tissue specific expression of retinoblastoma protein
JOURNAL Patent: JP 2001503638-A 3 21-MAR-2001;
COMMENT CANJI INC
OS Unidentified
PN JP 2001503638-A/3
PD 21-MAR-2001
PF 13-NOV-1997 JP 1998522958
PR 15-NOV-1996 US 08/751517,14-FEB-1997 US 08/801092 PI
DOUGLAS ANTELMAN,RICHARD J GREGORY,KENNETH N WILLS PC
C07H21/04,C07K5/00,A61K38/00,A61K35/12
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT CDS Location/Qualifiers
source 1..3853
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

FEATURES
ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 3853;
Best Local Similarity 100.0%; Pred. No. 3.3e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTTGTGTGTGGAGTCTGCTGAGTAGTGCAGCAAAATTAAGCTACA 60
Db 81 CTGCTCCCTGCTTGTGTGTGGAGTCTGCTGAGTAGTGCAGCAAAATTAAGCTACA 140

```

```

Qy 61 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 100
Db 141 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 180

```

```

RESULT 13
AR207832 LOCUS 3853 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 5 from patent US 6379927.
ACCESSION AR207832
VERSION AR207832.1 GI:21507688
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3853)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion proteins
JOURNAL Patent: US 6379927-A 5 30-APR-2002;
FEATURES Location/Qualifiers
source 1..3853
/organism="unknown"
/mol_type="unassigned DNA"

```

```

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 3853;
Best Local Similarity 100.0%; Pred. No. 3.3e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 CTGCTCCCTGCTTGTGTGTGGAGTCTGCTGAGTAGTGCAGCAAAATTAAGCTACA 60
Db 81 CTGCTCCCTGCTTGTGTGTGGAGTCTGCTGAGTAGTGCAGCAAAATTAAGCTACA 140

Qy 61 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 100
Db 141 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 180

```

```

RESULT 14
BD009729 LOCUS 3853 bp DNA linear PAT 31-JAN-2002
DEFINITION Tissue specific expression of retinoblastoma protein.
ACCESSION BD009729
VERSION BD009729.1 GI:18638102
KEYWORDS JP 2001503638-A/3.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 3853)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Tissue specific expression of retinoblastoma protein
JOURNAL Patent: JP 2001503638-A 3 21-MAR-2001;
COMMENT CANJI INC
OS Unidentified
PN JP 2001503638-A/3
PD 21-MAR-2001
PF 13-NOV-1997 JP 1998522958
PR 15-NOV-1996 US 08/751517,14-FEB-1997 US 08/801092 PI
DOUGLAS ANTELMAN,RICHARD J GREGORY,KENNETH N WILLS PC
C07H21/04,C07K5/00,A61K38/00,A61K35/12
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT CDS Location/Qualifiers
source 1..3853
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

```

```

FEATURES
ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 3853;
Best Local Similarity 100.0%; Pred. No. 3.3e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 CTGCTCCCTGCTTGTGTGTGGAGTCTGCTGAGTAGTGCAGCAAAATTAAGCTACA 60
Db 81 CTGCTCCCTGCTTGTGTGTGGAGTCTGCTGAGTAGTGCAGCAAAATTAAGCTACA 140

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:35:42 ; Search time 140.988 Seconds
(without alignments)
4198.742 Million cell updates/sec

Title: US-09-482-682-64_COPY_1_100
Perfect score: 100
Sequence: 1 ctgtccctctgtgtgtt.....caattgatgaagaattgc 100

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: geneseq1980s:*
- 2: geneseq1990s:*
- 3: geneseq2000s:*
- 4: geneseq2001as:*
- 5: geneseq2001bs:*
- 6: geneseq2002as:*
- 7: geneseq2002bs:*
- 8: geneseq2003as:*
- 9: geneseq2003bs:*
- 10: geneseq2003cs:*
- 11: geneseq2003ds:*
- 12: geneseq2004as:*
- 13: geneseq2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	100	100.0	562	8	ABZ23250 Nucleotid
2	100	100.0	648	4	AH43951 Rous sarc
3	100	100.0	1070	2	AAV58058 Plasmid C
4	100	100.0	1506	12	ADMA1035 Fungus nu
5	100	100.0	1600	2	ADH11349 Vertebrat
6	100	100.0	1782	12	ADMA1037 Cytomegal
7	100	100.0	2241	12	ADMA1034 Human nuc
8	100	100.0	2245	8	ABZ23249 Lac repre
9	100	100.0	2294	12	ADMA1036 Cytomegal
10	100	100.0	2426	4	AD02037 Plasmid p
11	100	100.0	2427	4	AD02036 Plasmid p
12	100	100.0	3400	2	AAT62937 3F4 human
13	100	100.0	3400	2	AAT62932 2A2 human
14	100	100.0	3853	2	AAV40006 Plasmid p
15	100	100.0	3925	2	AAT90695 Plasmid C
16	100	100.0	4026	2	AAV40007 Plasmid p
17	100	100.0	4059	2	AAQ75974 pHLA-B7 e
18	100	100.0	4249	2	AAV63466 Plasmid p
19	100	100.0	4341	2	AAQ62391 Vector pv
20	100	100.0	4341	6	AA517704 Vector pv

21	100	100.0	4341	6	ABN83143	Abn83143 Plasmid p
22	100	100.0	4457	10	ADD35599	Add35599 Bicistron
23	100	100.0	4525	2	AAV69746	Aav69746 Nucleotid
24	100	100.0	4597	4	AAF24901	Aaf24901 Nucleotid
25	100	100.0	4825	13	ADR12380	Adr12380 Vector pm
26	100	100.0	4840	4	AAF83146	Aaf83146 Complete
27	100	100.0	4965	2	AAQ75973	Aaq75973 pHLA-B7/b
28	100	100.0	5015	10	ADB33528	Adb33528 Expressio
29	100	100.0	5053	3	AAZ38633	Aaz38633 pEP2 expr
30	100	100.0	5070	4	AA512839	Aa512839 DNA sequ
31	100	100.0	5082	2	ADH11417	Adh11417 plasmid p
32	100	100.0	5162	10	ADF10526	Adf10526 Plasmid p
33	100	100.0	5162	10	ACC44637	Acc44637 Murine rd
34	100	100.0	5172	13	ADS75099	Ads75099 Plasmid p
35	100	100.0	5173	6	ABK88869	Abk88869 Topoisome
36	100	100.0	5173	12	ADB83792	Adb83792 Plasmid p
37	100	100.0	5173	12	ADO06721	Ado06721 Recombina
38	100	100.0	5192	10	ACC44692	Acc44692 Plasmid p
39	100	100.0	5250	2	AAT62933	Aat62933 2A2 human
40	100	100.0	5271	10	ABV77540	Abv77540 Plasmid p
41	100	100.0	5283	10	ABV77538	Abv77538 Plasmid p
42	100	100.0	5292	10	ABV77547	Abv77547 Plasmid p
43	100	100.0	5293	10	ABV77548	Abv77548 Plasmid p
44	100	100.0	5293	10	ABV77549	Abv77549 Plasmid p
45	100	100.0	5300	2	AAT62938	Aat62938 3F4 human

ALIGNMENTS

RESULT 1
ABZ23250
ID ABZ23250 standard; DNA; 562 BP.
AC ABZ23250;
XX
DT 24-MAR-2003 (first entry)
XX
DB Nucleotide sequence of the Rous sarcoma virus (RSV)-LTR promoter.
KW p21; RSV; LTR promoter; cell cycle inhibitor protein; protein production;
KW anchorage-independent producer cell line; ss.
XX
OS Rous sarcoma virus.
XX
PN WO200299100-A2.
XX
PD 12-DEC-2002.
XX
PF 03-JUN-2002; 2002WO-EP006054.
XX
PR 01-JUN-2001; 2001GB-00013318.
XX
PA (LONZ) LONZA BIOLOGICS PLC.
PI Al-Rubeai M, Shuttleworth J;
XX
DR WPI; 2003-148669/14.
XX
PT Producing recombinant protein, particularly for maximizing or enhancing
PT e.g. therapeutic protein production, by co-expressing protein with
PT recombinant cell cycle inhibitor protein (p21) in producer cell line.
XX
FS Disclosure; Page 32-33; 33pp; English.
XX
CC The present sequence represents the Rous sarcoma virus (RSV)-LTR
CC promoter. The present sequence is used to produce vectors for use in the
CC method of the invention. The specification describes a method for
CC producing a protein, preferably a recombinant protein, in a mammalian
CC anchorage-independent producer cell line. The method comprises co-
CC expressing with the protein in the producer cell line a recombinant cell
CC cycle inhibitor protein (preferably p21). The method is useful for
CC producing a recombinant protein in a producer cell line. This is

CC particularly useful for maximizing or enhancing the production of e.g.
CC therapeutic proteins at an industrial scale
SQ Sequence 562 BP; 143 A; 109 C; 163 G; 147 T; 0 U; 0 Other;
Query Match 100.0%; Score 100; DB 8; Length 562;
Best Local Similarity 100.0%; Pred. No. 8.2e-28;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTGCTCCCTGTTGTGTTGGAGTGCCTGAGTAGTGCAGCAAAATTTAAGCTACA 60
Db 46 CTGCTCCCTGTTGTGTTGGAGTGCCTGAGTAGTGCAGCAAAATTTAAGCTACA 105
Qy 61 ACAAGGCAAGCTTGACCGACAATTGCATGAAGAATCTGC 100
Db 106 ACAAGGCAAGCTTGACCGACAATTGCATGAAGAATCTGC 145
RESULT 2
AAH43951
ID AAH43951 standard; DNA; 648 BP.
XX AC AAH43951;
XX DT 06-SEP-2001 (first entry)
XX DE Rous sarcoma virus promoter nucleotide sequence SEQ ID NO:1.
XX KW Rous sarcoma virus; promoter; enhancer; RSV; primate; gene expression;
XX KW transgene; genetic engineering; gene therapy; immunisation; ds.
XX OS Rous sarcoma virus.
XX FN WO200142444-A2.
XX PD 14-JUN-2001.
XX PF 08-DEC-2000; 2000WO-US033256.
XX PR 10-DEC-1999; 99US-0170019P.
XX PA (ARIA-) ARIAD GENE THERAPEUTICS INC.
XX PA (UYPE-) UNIV PENNSYLVANIA.
XX PI Rivera V, Zoltick P, Wilson JM;
XX WPI; 2001-381673/40.
Genetically engineering a primate for expression of a desired gene,
comprises introducing into the primate a transgene comprising Rous
Sarcoma Virus (RSV) promoter and a nucleic acid sequence heterologous to
RSV promoter.
Claim 7; Page 44; 64pp; English.
The present invention describes a method for genetically engineering a
primate for expression of a desired gene comprising introducing into the
primate a transgene comprising an Rous Sarcoma Virus (RSV) promoter and a
nucleic acid sequence heterologous to RSV promoter. Also described is a
primate cell (I) containing and capable of expressing a transgene
comprising an RSV promoter operably linked to a recombinant nucleic acid
encoding one or more fusion proteins, where the fusion proteins bind to a
ligand and in the presence of the ligand modulate(s) the expression level
of a target gene. The method can be used for high level expression of
genes in primates or for engineering primate cells. It is useful for
increasing the efficacy of many gene therapy strategies, and for
increasing the efficacy of intracellular immunisation agents, molecules
like ribozymes, antisense RNA, and dominant negative proteins, that act
either stoichiometrically, or by competition. The method increases the
efficacy of many gene therapy strategies by substantially elevating the
expression of an exogenous therapeutic gene, and allowing expression to
reach therapeutically effective levels. The present sequence represents a
specifically claimed RSV enhancer/promoter nucleotide sequence from the

CC present invention
XX SQ Sequence 648 BP; 163 A; 135 C; 179 G; 171 T; 0 U; 0 Other;
Query Match 100.0%; Score 100; DB 4; Length 648;
Best Local Similarity 100.0%; Pred. No. 8.6e-28;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTGCTCCCTGTTGTGTTGGAGTGCCTGAGTAGTGCAGCAAAATTTAAGCTACA 60
Db 90 CTGCTCCCTGTTGTGTTGGAGTGCCTGAGTAGTGCAGCAAAATTTAAGCTACA 149
Qy 61 ACAAGGCAAGCTTGACCGACAATTGCATGAAGAATCTGC 100
Db 150 ACAAGGCAAGCTTGACCGACAATTGCATGAAGAATCTGC 189
RESULT 3
AAV58058
ID AAV58058 standard; DNA; 1070 BP.
XX AC AAV58058;
XX DT 27-AUG-2003 (revised)
XX DT 11-JAN-1999 (first entry)
XX DE Plasmid CMV-delRT SstI fragment.
XX KW FIV; FIPV; vaccine; reverse transcriptase; diagnosis; therapy; CMV-delRT;
XX KW promoter; cat; ss.
XX OS Human cytomegalovirus.
XX OS feline immunodeficiency virus.
XX OS Chimeric.
XX FH Key Location/Qualifiers
FT Promoter 8...896
FT /*tag= a
FT /note= "CMV promoter fragment from pcDNA3 (BgIII-KpnII)"
FT provirus 918..1070
FT /*tag= b
FT /note= "FIV sequences from primer binding site to SstI
FT site"
XX W09840493-A1.
XX 17-SEP-1998.
XX 10-MAR-1998; 98WO-GB000715.
XX 11-MAR-1997; 97GB-00004977.
XX (UNIU) UNIV GLASGOW.
XX Neil JC, Rigby MA, Jarrett JO;
XX WPI; 1998-520813/44.
Protecting, e.g. cats, against feline immunodeficiency virus - by using
vaccine comprising FIV pol gene containing deletion and/or insertion in
reverse transcriptase domain.
Example 3; Fig 4; 66pp; English.
This is the nucleotide sequence of a SstI fragment of plasmid CMV-delRT,
in which the immediate-early promoter of human cytomegalovirus replaces
the 5' long terminal repeat region of feline immunodeficiency virus (FIV)
clone F14-delRT (see AAV58053). FIV sequences downstream of the SstI site
are identical to those in F14-delRT. Use of the CMV promoter was designed
to enhance expression of FIV antigens, and to reduce the risk of
reversion to a replicating provirus, in tissues after inoculation of DNA.
Vaccine formulations for FIV-related diseases include a defective feline
immunodeficiency proviral (FIPV) polynucleotide comprising an in-frame

CC deletion and/or insertion in the reverse transcriptase (RT) region of the
CC pol gene. Host cells comprising the PIPV are capable of producing FIV
CC proteins, except for functionally competent RT, and thus release non-
CC infectious FIV viral particles. (Updated on 27-AUG-2003 to correct OS
CC field.)

XX Sequence 1070 BP; 275 A; 254 C; 268 G; 273 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 2; Length 1070;

Best Local Similarity 100.0%; Pred. No. 1e-27;

Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTGTGTGGAGTCTGCTGAGTAGTGGCGAGCAAAATTTAAGCTACA 60

Db 77 CTGCTCCCTGCTGTGTGGAGTCTGCTGAGTAGTGGCGAGCAAAATTTAAGCTACA 136

Qy 61 ACAAGGCAAGCTTGACCGACAATTGCAATGCAATGCAATCTGC 100

Db 137 ACAAGGCAAGCTTGACCGACAATTGCAATGCAATGCAATCTGC 176

RESULT 4

ADM41035

ID ADM41035 standard; DNA; 1506 BP.

XX

AC ADM41035;

XX

DT 17-JUN-2004 (first entry)

XX

DE Fungus nucleotide sequence SEQ ID NO:3.

XX

KW engrafting foreign replacement cell; implanting foreign replacement cell;

KW growth; differentiation; drug development; vaccine development;

KW tissue transplantation; human disease study; fungus; gene; ds.

XX

OS Unidentified.

XX

PN WO2004027029-A2.

XX

PD 01-APR-2004.

XX

PF 17-SEP-2003; 2003WO-US029251.

XX

PR 19-SEP-2002; 2002US-0411790P.

XX

PA (XIME-) XIMEREX INC.

XX

PI Beschornier WE, Sosa CE, Thompson SC;

XX

DR WPI; 2004-295402/27.

XX

PS Engrafting foreign replacement cells within a fetal non-human mammal,

XX

CC useful in producing chimeric mammals, comprises selectively destroying

CC native cells in a tissue of a fetal non-human mammal host.

XX

PS Disclosure; SEQ ID NO 3; 48pp; English.

XX

CC The present invention describes a method for engrafting foreign

CC replacement cells within a foetal non-human mammal, which comprises

CC selectively destroying native cells in a tissue of a foetal non-human

CC mammal host, where the number of maternal cells of the same tissue is not

CC substantially reduced, and implanting foreign replacement cells in the

CC tissue of the fetal non-human mammal host, where the foreign replacement

CC cells replace destroyed cells of the tissue. The method is useful for

CC facilitating growth and differentiation of foreign cells within a

CC mammalian host, and for producing chimeric mammals that can be used to

CC develop new drugs and vaccine, factors, drugs and tissues for

CC transplantation, also useful to study human diseases. The present

CC sequence represents a nucleotide sequence given in the Sequence Listing

CC of the present invention but not mentioned further within the

XX specification.

XX Sequence 1506 BP; 454 A; 277 C; 361 G; 414 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 12; Length 1506;

Best Local Similarity 100.0%; Pred. No. 1.2e-27;

Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTGTGTGGAGTCTGCTGAGTAGTGGCGAGCAAAATTTAAGCTACA 60

Db 81 CTGCTCCCTGCTGTGTGGAGTCTGCTGAGTAGTGGCGAGCAAAATTTAAGCTACA 140

Qy 61 ACAAGGCAAGCTTGACCGACAATTGCAATGCAATGCAATCTGC 100

Db 141 ACAAGGCAAGCTTGACCGACAATTGCAATGCAATGCAATCTGC 180

RESULT 5

ADH11349

ID ADH11349 standard; DNA; 1600 BP.

XX

AC ADH11349;

XX

DT 11-MAR-2004 (first entry)

XX

DE Vertebrate UNC-53 protein homologue related nucleotide sequence.

XX

KW UNC-53 vertebrate protein homologue; UNC-53; Caenorhabditis elegans;

KW cell shape regulator; cell motility regulator; cell migration;

KW cell behaviour regulator; phenotype; signal transduction pathway;

KW signal transducing protein; signal integrator protein;

KW neuronal regeneration; revascularisation; wound healing;

KW chronic neurodegenerative disease; acute traumatic injury;

KW fibrotic disease; gene; ds.

XX

OS Unidentified.

XX

FN WO9824810-A2.

XX

PD 11-JUN-1998.

XX

PF 03-DEC-1997; 97WO-EP006956.

XX

PR 04-DEC-1996; 96GB-00025283.

XX

PA (JANC) JANSSEN PHARM NV.

XX

PI Platteeuw CJ, Buesa Arjol CM, Deraeymaeker M, Verhasselt P;

XX

PI Pujol NJR, Maertens LJS, Luyten W, Geerts H, Vandekerckhove JS;

XX

PI Geysen J, Bogaert TA0E;

XX

DR WPI; 1998-362411/31.

XX

DR P-PSDB; ADH11350.

XX

PT Vertebrate homologues of C. elegans UNC-53 protein - useful for, e.g.

XX

PT promoting neuronal regeneration, treating chronic neuro-degenerative

XX

PT diseases or acute traumatic injuries.

XX

PS Disclosure; Page 410-411; 479pp; English.

XX

CC The present invention describes a vertebrate protein homologue of an UNC-

CC 53 protein of Caenorhabditis elegans or a functional equivalent,

CC derivative or bioprecursor of UNC-53. Also described: (1) a cDNA sequence

CC encoding a vertebrate homologue of the C. elegans UNC-53 protein; (2) a

CC nucleic acid which hybridises to the cDNA of (1); (3) vector comprising

CC the cDNA as in (1); (4) a host cell containing the vector as in (3); (5)

CC a transgenic cell, tissue or animal comprising the vector as in (3); (6)

CC a compound identified as an enhancer or inhibitor of the regulation of

CC cell shape, motility, or the direction of cell migration for use as a

CC therapeutic; (7) a method for determination of whether a protein is an

CC inhibitor or enhancer of regulation of cell behaviour, growth, shape or

CC motility or the direction of migration by contacting a host cell

CC expressing a homologue of UNC-53 and determining a change of phenotype;

CC (8) a method for identification of vertebrate homologues of C. elegans

CC unc-53 gene by hybridising a probe of 15-50 bp of the unc-53 sequence to

CC a DNA library; and (9) a method for identification of a protein which is

CC active in the signal transduction pathway of a cell of which a vertebrate
CC homologue of UNC-53 is a component comprising: (i) contacting an extract
CC of a cell with an antibody to the UNC-53 homologue; (ii) identifying an
CC antibody/homologue complex; and (iii) analysing such a complex to
CC identify any non-antibody protein bound to the complex. UNC-53 is a
CC signal transducing or signal integrator protein involved in controlling
CC directionality of cell migration and cell shape in *C. elegans*. Vertebrate
CC homologues of UNC-53 can be used to promote neuronal regeneration,
CC revascularisation or wound healing, to treat chronic neurodegenerative
CC diseases or acute traumatic injuries or fibrotic diseases. The present
CC sequence is used in the exemplification of the present invention.

SQ Sequence 1600 BP; 397 A; 409 C; 398 G; 396 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 2; Length 1600;
Best Local Similarity 100.0%; Pred. No. 1.2e-27;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCTCCCTGCTGTGTGGAGTGCCTGAGTAGTGGCGAGCAAAATTTAAGCTTACA 60
DB 81 CTGCTCCCTGCTGTGTGGAGTGCCTGAGTAGTGGCGAGCAAAATTTAAGCTTACA 140

QY 61 ACAAGGCAAGGCTTGACCGCAAAATTCATGAAGAATCTGC 100
DB 141 ACAAGGCAAGGCTTGACCGCAAAATTCATGAAGAATCTGC 180

RESULT 6
ADM41037
ID ADM41037 standard; DNA; 1782 BP.
AC ADM41037;
XX
XX
XX 17-JUN-2004 (first entry)
XX
XX
XX Cytomegalovirus nucleotide sequence SEQ ID NO:5.
XX
XX engrafting foreign replacement cell; implanting foreign replacement cell;
XX growth; differentiation; drug development; vaccine development;
XX tissue transplantation; human disease study; cytomegalovirus; gene; ds.
XX
XX Cytomegalovirus.
XX
XX WO2004027029-A2.
XX
XX 01-APR-2004.
XX
XX 17-SEP-2003; 2003WO-US029251.
XX
XX 19-SEP-2002; 2002US-0411790P.
XX (XIME-) XIMEREX INC.
XX
XX Beschorner WE, Sosa CE, Thompson SC;
XX WPI; 2004-295402/27.
XX
XX Engrafting foreign replacement cells within a fetal non-human mammal,
XX useful in producing chimeric mammals, comprises selectively destroying
XX native cells in a tissue of a fetal non-human mammal host.

PS Disclosure; SEQ ID NO 5; 48pp; English.

CC The present invention describes a method for engrafting foreign
CC replacement cells within a fetal non-human mammal, which comprises
CC selectively destroying native cells in a tissue of a fetal non-human
CC mammal host, where the number of maternal cells of the same tissue is not
CC substantially reduced, and implanting foreign replacement cells in the
CC tissue of the fetal non-human mammal host, where the foreign replacement
CC cells replace destroyed cells of the tissue. The method is useful for
CC facilitating growth and differentiation of foreign cells within a
CC mammalian host, and for producing chimeric mammals that can be used to
CC develop new drugs and vaccine, factors, drugs and tissues for

CC transplantation, also useful to study human diseases. The present
CC sequence represents a nucleotide sequence given in the Sequence Listing
CC of the present invention but not mentioned further within the
CC specification.

SQ Sequence 1782 BP; 458 A; 399 C; 451 G; 474 T; 0 U; 0 Other;
Query Match 100.0%; Score 100; DB 12; Length 1782;
Best Local Similarity 100.0%; Pred. No. 1.2e-27;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCTCCCTGCTGTGTGGAGTGCCTGAGTAGTGGCGAGCAAAATTTAAGCTTACA 60
DB 81 CTGCTCCCTGCTGTGTGGAGTGCCTGAGTAGTGGCGAGCAAAATTTAAGCTTACA 140

QY 61 ACAAGGCAAGGCTTGACCGCAAAATTCATGAAGAATCTGC 100
DB 141 ACAAGGCAAGGCTTGACCGCAAAATTCATGAAGAATCTGC 180

RESULT 7
ADM41034
ID ADM41034 standard; DNA; 2241 BP.
AC ADM41034;
XX
XX 17-JUN-2004 (first entry)
XX
XX Human nucleotide sequence SEQ ID NO:2.
XX
XX engrafting foreign replacement cell; implanting foreign replacement cell;
XX growth; differentiation; drug development; vaccine development;
XX tissue transplantation; human disease study; human; gene; ds.
XX
XX Homo sapiens.
XX
XX WO2004027029-A2.
XX
XX 01-APR-2004.
XX
XX 17-SEP-2003; 2003WO-US029251.
XX
XX 19-SEP-2002; 2002US-0411790P.
XX (XIME-) XIMEREX INC.
XX
XX Beschorner WE, Sosa CE, Thompson SC;
XX WPI; 2004-295402/27.

XX Engrafting foreign replacement cells within a fetal non-human mammal,
XX useful in producing chimeric mammals, comprises selectively destroying
XX native cells in a tissue of a fetal non-human mammal host.

PS Disclosure; SEQ ID NO 2; 48pp; English.

CC The present invention describes a method for engrafting foreign
CC replacement cells within a fetal non-human mammal, which comprises
CC selectively destroying native cells in a tissue of a fetal non-human
CC mammal host, where the number of maternal cells of the same tissue is not
CC substantially reduced, and implanting foreign replacement cells in the
CC tissue of the fetal non-human mammal host, where the foreign replacement
CC cells replace destroyed cells of the tissue. The method is useful for
CC facilitating growth and differentiation of foreign cells within a
CC mammalian host, and for producing chimeric mammals that can be used to
CC develop new drugs and vaccine, factors, drugs and tissues for
CC transplantation, also useful to study human diseases. The present
CC sequence represents a nucleotide sequence given in the Sequence Listing
CC of the present invention but not mentioned further within the
CC specification.

SQ Sequence 2241 BP; 498 A; 630 C; 609 G; 504 T; 0 U; 0 Other;

Qy 1 CTGCTCCCTGCTTGTGTGAGGTCGCTGAGTAGTGGCGAGCAAAATTTAGCTACA 60
 |||||
 Db 81 CTGCTCCCTGCTTGTGTGAGGTCGCTGAGTAGTGGCGAGCAAAATTTAGCTACA 140
 |||||
 Qy 61 ACAAGGCAAGGCTTGACCGACAATTGTCATGAAGAATCTGC 100
 |||||
 Db 141 ACAAGGCAAGGCTTGACCGACAATTGTCATGAAGAATCTGC 180
 |||||

RESULT 10
 AAD02037
 ID AAD02037 standard; DNA; 2426 BP.
 XX
 AC AAD02037;
 XX
 DT 11-SEP-2003 (revised)
 DT 26-MAR-2001 (first entry)
 XX
 XX Plasmid pNG3/RC/CPG2-Thy1 comprising CPG2 DNA with rat thy1 gene.
 XX Carboxypeptidase G2; CPG2; gene directed enzyme prodrug therapy; GDEPT;
 KW glycosylphosphatidylinositol; GPI; cancer; therapy; rat; Thy1 gene;
 KW plasmid; ds.
 XX Rattus sp.
 OS Bacteria.
 OS Chimeric.
 XX WO200066752-A2.
 XX 09-NOV-2000.
 XX 28-APR-2000; 2000WO-GB001640.
 XX 01-MAY-1999; 99GB-00010077.
 XX (ASTR) ASTRAZENECA AB.
 PA (UTMA-) UNIV VICTORIA MANCHESTER.
 XX Castro MG, Emery SC, Lowenstein PR;
 PI WPI; 2001-015983/02.
 XX
 XX Gene directed enzyme prodrug therapy using post translational
 PT glycosylphosphatidylinositol addition to prodrug activating enzyme to
 PT enable anchorage of enzyme at cell surface for cancer therapy.
 XX
 PS Example 1e; Page 59-60; 60pp; English.
 XX
 XX The present invention relates to a gene directed enzyme prodrug therapy
 CC (GDEPT) using post translational glycosylphosphatidylinositol (GPI)
 CC addition to a prodrug activating enzyme which enables anchorage of the
 CC enzyme at the cell surface. Carboxypeptidase G2 (CPG2) is a preferred
 CC prodrug activating enzyme. The invention also relates to an expression
 CC vector for expression of a GPI enzyme hybrid capable of anchorage to the
 CC surface of a mammalian cell. The expression vector comprise
 CC polynucleotide sequences encoding a signal peptide, an enzyme capable of
 CC activating a prodrug, and a post-translational GPI addition motif. The
 CC expression vector is useful in the manufacture of a medicament for cancer
 CC therapy in a mammalian host. The present DNA sequence is a plasmid
 CC pNG3/RC/CPG2-Thy1 comprising CPG2 nucleic acid sequence with the last
 CC exon of rat thy 1 gene at its 3' end. (Updated on 11-SEP-2003 to
 CC standardise OS field)
 XX
 SQ Sequence 2426 BP; 557 A; 705 C; 668 G; 495 T; 0 U; 1 Other;

Query Match 100.0%; Score 100; DB 4; Length 2426;
 Best Local Similarity 100.0%; Pred. No. 1.4e-27;
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTTGTGTGAGGTCGCTGAGTAGTGGCGAGCAAAATTTAGCTACA 60
 |||||
 Db 69 CTGCTCCCTGCTTGTGTGAGGTCGCTGAGTAGTGGCGAGCAAAATTTAGCTACA 128
 |||||

Qy 61 ACAAGGCAAGGCTTGACCGACAATTGTCATGAAGAATCTGC 100
 |||||
 Db 129 ACAAGGCAAGGCTTGACCGACAATTGTCATGAAGAATCTGC 168
 |||||

RESULT 11
 AAD02036
 ID AAD02036 standard; DNA; 2427 BP.
 XX
 AC AAD02036;
 XX
 DT 11-SEP-2003 (revised)
 DT 26-MAR-2001 (first entry)
 XX
 XX Plasmid pNG3/RC/CPG2(Q3)-Thy1 comprising CPG2 variant with rat thy1 gene.
 XX Carboxypeptidase G2; CPG2; gene directed enzyme prodrug therapy; GDEPT;
 KW glycosylphosphatidylinositol; GPI; cancer; therapy; rat; Thy1 gene;
 KW CPG2(Q3) variant; plasmid; ds.
 XX Rattus sp.
 OS Bacteria.
 OS Chimeric.
 XX WO200066752-A2.
 XX 09-NOV-2000.
 XX 28-APR-2000; 2000WO-GB001640.
 XX 01-MAY-1999; 99GB-00010077.
 XX (ASTR) ASTRAZENECA AB.
 PA (UTMA-) UNIV VICTORIA MANCHESTER.
 XX Castro MG, Emery SC, Lowenstein PR;
 PI WPI; 2001-015983/02.
 XX
 XX Gene directed enzyme prodrug therapy using post translational
 PT glycosylphosphatidylinositol addition to prodrug activating enzyme to
 PT enable anchorage of enzyme at cell surface for cancer therapy.
 XX
 PS Example 1e; Page 59; 60pp; English.
 XX
 XX The present invention relates to a gene directed enzyme prodrug therapy
 CC (GDEPT) using post translational glycosylphosphatidylinositol (GPI)
 CC addition to a prodrug activating enzyme which enables anchorage of the
 CC enzyme at the cell surface. Carboxypeptidase G2 (CPG2) is a preferred
 CC prodrug activating enzyme. The invention also relates to an expression
 CC vector for expression of a GPI enzyme hybrid capable of anchorage to the
 CC surface of a mammalian cell. The expression vector comprise
 CC polynucleotide sequences encoding a signal peptide, an enzyme capable of
 CC activating a prodrug, and a post-translational GPI addition motif. The
 CC expression vector is useful in the manufacture of a medicament for cancer
 CC therapy in a mammalian host. The present DNA sequence is a plasmid
 CC pNG3/RC/CPG2(Q3) comprising CPG2 variant CPG2(Q3) and the last exon of
 CC rat Thy-1 at the 3' end. (Updated on 11-SEP-2003 to standardise OS field)
 XX
 SQ Sequence 2427 BP; 555 A; 706 C; 670 G; 495 T; 0 U; 1 Other;

Query Match 100.0%; Score 100; DB 4; Length 2427;
 Best Local Similarity 100.0%; Pred. No. 1.4e-27;
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTTGTGTGAGGTCGCTGAGTAGTGGCGAGCAAAATTTAGCTACA 60
 |||||
 Db 70 CTGCTCCCTGCTTGTGTGAGGTCGCTGAGTAGTGGCGAGCAAAATTTAGCTACA 129
 |||||
 Qy 61 ACAAGGCAAGGCTTGACCGACAATTGTCATGAAGAATCTGC 100
 |||||
 Db 130 ACAAGGCAAGGCTTGACCGACAATTGTCATGAAGAATCTGC 169
 |||||

PT Selectable retroviral packaging cell lines and expression constructs -
PT comprise selectable gene downstream of gene of interest, are selectable
PT due to the in-efficiency associated with translation re-initiation.
XX
PS
PS Claim 23; Fig 13; 79pp; English.
XX
CC This sequence represents the recombinant expression plasmid CMV10A. This
CC sequence is a packaging-deficient construct having a viral env gene (in
CC this case from moloney murine leukaemia virus under hCMV promoter
CC control) and a selectable marker (SM). It is an example of a recombinant
CC expression vector (REV) of the invention, used to create a packaging cell
CC line. The REV's of the invention comprise a gene of interest (GOI) and a
CC SM gene. The SM gene is arranged downstream of the GOI and a GOI
CC associated stop codon is spaced from a start codon of the SM gene to
CC ensure that the SM protein is expressed as a result of translation
CC reinitiation. The cell lines are transformed with two REV's, both are
CC replication deficient, one contains the viral gag-pol gene, the other the
CC viral env gene. By using helper constructs, such as the REV's, which are
CC directly selectable and which provide for high expression of the viral
CC gene, high titre retroviral vectors may be obtained. The packaging cell
CC lines are useful for gene therapy. Prior packaging cell lines using full
CC length retroviral genomes as helper genomes were isolated by
CC cotransfecting them with plasmids encoding selectable markers. However,
CC the helper functions can be lost during the passages of the cells in
CC culture and the current packaging systems provide limited titres of
CC infectious retroviral vectors. Co-transfection with a plasmid encoding a
CC SM does not directly select the best gag-pol-env-expressing cells. The
CC new retroviral packaging cell lines overcome these problems
XX
SQ Sequence 3925 BP; 963 A; 1001 C; 959 G; 998 T; 0 U; 4 Other;

Query Match 100.0%; Score 100; DB 2; Length 3925;
Best Local Similarity 100.0%; Pred. No. 1.6e-27;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTGCTCCCTGCTTGTGTGGAGTGCCTGAGTAGTGGCGGAGCAAAATTTAGCTACA 60
Db 70 CTGCTCCCTGCTTGTGTGGAGTGCCTGAGTAGTGGCGGAGCAAAATTTAGCTACA 129
Qy 61 ACAAGGCAAGGCTTGACCGACAATTGCATGAGAAATCTGC 100
Db 130 ACAAGGCAAGGCTTGACCGACAATTGCATGAGAAATCTGC 169

Search completed: July 14, 2005, 07:01:45
Job time : 143.038 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 05:15:57 ; Search time 952.146 Seconds
(without alignments)
3997.736 Million cell updates/sec

Title: US-09-482-682-64_COPY_1_100

Perfect score: 100

Sequence: 1 ctgtccctcgttgtgtgtt.....caattgcatgaagaattgc 100

Scoring table:

IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	100	100.0	602	8	B67169 CPG0047A Cp
2	30.4	30.4	829	4	BI333630 602997459
3	30.2	30.2	823	6	CD655614 AGENCOURT
4	29.8	29.8	1165	8	CC242469 CH261-11F
5	29.4	29.4	401	6	CB387202 OSTF076E6
6	29.4	29.4	531	5	BQ310441 MR0-BT450
7	29.4	29.4	754	8	AQ946479 Sheared D
8	29.2	29.2	426	9	CC888514 SALK 1519
9	29	29.0	657	7	CK086063 RG11-C07
10	29	29.0	877	9	AL225351 Tetraodon
11	28.6	28.6	340	7	F32722 HSPD25699 H
12	28.6	28.6	408	1	AA962465 oc91e05.s
13	28.6	28.6	436	7	CN386744 328755673
14	28.6	28.6	514	6	CB161201 K-EST0221
15	28.6	28.6	530	6	CB161182 K-EST0220
16	28.6	28.6	534	2	AW500392 UI-HF-BN0
17	28.6	28.6	550	1	AA984313 am83h04.s
18	28.6	28.6	555	7	CR537056 DKFZp459D
19	28.6	28.6	583	7	CN386728 170005326
20	28.6	28.6	585	4	BG993413 MR3-HT099
21	28.6	28.6	625	4	BI113747 602860946
22	28.6	28.6	626	7	CN386724 170006000
23	28.6	28.6	641	2	AW955076 EST367146
24	28.6	28.6	651	7	CN386730 170005316

c 25 28.6 28.6 680 7 CN386731 170005313
 c 26 28.6 28.6 682 7 CN386746 170005999
 c 27 28.6 28.6 687 7 CN386717 170004551
 c 28 28.6 28.6 710 2 BF309673 601891808
 c 29 28.6 28.6 724 7 CN386780 170005318
 c 30 28.6 28.6 770 6 CD654097 AGENCOURT
 c 31 28.6 28.6 777 7 CN386742 170005314
 c 32 28.6 28.6 801 6 CD656906 AGENCOURT
 c 33 28.6 28.6 814 4 BG820298 602782110
 c 34 28.6 28.6 823 6 CB996419 AGENCOURT
 c 35 28.6 28.6 827 4 BG387788 602412672
 c 36 28.6 28.6 830 5 CD643822 AGENCOURT
 c 37 28.6 28.6 842 5 BU170446 AGENCOURT
 c 38 28.6 28.6 851 6 CB993607 AGENCOURT
 c 39 28.6 28.6 852 5 BQ222104 AGENCOURT
 c 40 28.6 28.6 856 6 CD656102 AGENCOURT
 c 41 28.6 28.6 909 5 BU189860 AGENCOURT
 c 42 28.6 28.6 1025 4 BM477450 AGENCOURT
 c 43 28.4 28.4 579 6 CB239722 RSH15G08
 c 44 28.4 28.4 733 7 CF667137 RTCTNT1_28
 c 45 28.4 28.4 1015 6 BY703355 BY703355

ALIGNMENTS

RESULT 1
 B67169
 LOCUS B67169 602 bp DNA linear GSS 12-MAY-2000
 DEFINITION CPG0047A CpioWagDNA2 Cryptosporidium parvum genomic, genomic survey sequence.
 ACCESSION B67169
 VERSION B67169.1 GI:2642750
 KEYWORDS GSS.
 SOURCE Cryptosporidium parvum
 ORGANISM Cryptosporidium parvum
 Cryptosporidium parvum
 Rukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
 Cryptosporidiidae; Cryptosporidium.
 REFERENCE 1 (bases 1 to 602)
 AUTHORS Strong, W.B. and Nelson, R.G.
 TITLE Preliminary profile of the Cryptosporidium parvum genome: an expressed sequence tag and genome survey sequence analysis
 JOURNAL Mol. Biochem. Parasitol. 107 (1), 1-32 (2000)
 MEDLINE 20183851
 PUBMED 10717299
 COMMENT Contact: Nelson, R. G.
 Depts. of Medicine & Pharmaceutical Chemistry
 San Francisco General Hospital-University of California, San Francisco
 Box 0811, San Francisco, CA 94143-0811, USA
 Tel: 415 206 8846
 Fax: 415 206 3353
 Email: malariad@itsa.ucsf.edu
 Submitted sequence has been edited to remove vector sequences 5' to the insert, to correct miscalled bases and assign uncalled (N) bases throughout the sequence, and to terminate when base-calling became ambiguous.
 Seq primer: T7
 Class: Shotgun
 High quality sequence stop: 602.
 Location/Qualifiers
 1..602
 /organism="Cryptosporidium parvum"
 /mol_type="genomic DNA"
 /strain="IOWA"
 /db_xref="taxon:5807"
 /lab_host="E. coli XL2 Blue MRF"
 /clone_lib="CpioWagDNA2"
 /note="Vector: pCR-Script Amp SK+; Site 1: SrfI; C. parvum (IOWA isolate) genomic DNA was hydrodynamically sheared to produce fragments having a tight size distribution between 2-4 kb by Dr. Yvonne Thorstenson of the Stanford DNA Sequencing and Technology Center

(http://sequence-www.stanford.edu/group/techdev/shear.htm)

The randomly sheared gDNA was chromatographed on Sephacryl S-400 to remove any small fragments and DNA eluting in the void volume was blunt-ended with T4 DNA Polymerase, treated with alkaline phosphatase to prevent the ligation of multiple fragments, ligated to Srf I-digested PCR-Script Amp (SK+) vector and transformed into E. coli strain XL10 Gold. Recombinant clones from the first plating of the library were selected for sequence analysis using T3 and T7 primers."

ORIGIN

Query Match 100.0%; Score 100; DB 8; Length 602;
 Best Local Similarity 100.0%; Pred. No. 8e-24;
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGTCCTCGTCTGTGTGGAGTGCCTGAGTAGTGGCGAGCAAAATTTAAGCTACA 60
 |||||
 Db 41 CTGTCCTCGTCTGTGTGGAGTGCCTGAGTAGTGGCGAGCAAAATTTAAGCTACA 100
 |||||

Qy 61 ACAAGGCAAGCTTGACCGACAATTCGATGAAGAACTCTGC 100
 |||||
 Db 101 ACAAGGCAAGCTTGACCGACAATTCGATGAAGAACTCTGC 140
 |||||

RESULT 2

BI333630/c
 LOCUS
 DEFINITION 602997459P1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5139470 5',
 mRNA sequence.

ACCESSION BI333630
 VERSION BI333630.1 GI:15018287
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 829)
 NIH-MGC http://mgs.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: Incyte Genomics, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1AM11343 row: f column: 15
 High quality sequence stop: 788.
 Location/Qualifiers

FEATURES

source
 1..829
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5139470"
 /tissue_type="cervical carcinoma cell line"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_12"
 /note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.4 kb. Library prepared by Life
 Technologies."

ORIGIN

Query Match 30.4%; Score 30.4; DB 4; Length 829;
 Best Local Similarity 67.2%; Pred. No. 15;
 Matches 43; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy 19 TTGGAGGTCGTGAGTAGTGGCGAGCAAAATTTAAGCTACAACAGCAAGGCTTGACC 78
 |||||

Db 217 TTGGCGCTCCAGAAATTGTTGGTGAGCAAACTTCAAGTTGCTGCTGGGAGTCTTGACT 158
 Qy 79 GACA 82
 |||||
 Db 157 GACA 154

RESULT 3
 CD655614/c
 LOCUS
 DEFINITION CD655614 823 bp mRNA linear EST 18-JUN-2003
 (Long) Homo sapiens cDNA clone IMAGE:30424285 5', mRNA sequence.

ACCESSION CD655614
 VERSION CD655614.1 GI:31896113
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 823)
 NIH-MGC http://mgs.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Irene Ginis and Mahendra Rao, NIA
 cDNA Library Preparation: Yulan Piao and Minoru Ko
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC c lone distribution information
 can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: NDAM506 row: k column: 14
 High quality sequence stop: 640.
 Location/Qualifiers

FEATURES

source
 1..823
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30424285"
 /tissue_type="Embryonic Stem cells"
 /cell_line="WA01"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NIA Human H1 Embryonic Stem Cell cDNA Library
 (Long)"
 /note="Vector: pCMV-Sport6; Site 1: NotI; Site 2: SalI;
 This is a long-transcript enriched cDNA library (Genome
 Res. 11: 1553-1558 (2001). [PMID: 11544199]) from WA01
 cell line. Undifferentiated human ES cell line WA01/H1
 was obtained from WiCell Research Institute, Inc.,
 Madison, WI, cultured according to their instructions, on
 MEF feeders. They formed round colonies with defined edges
 and were positive for alkaline phosphatase, SSEA-4, OCT3,
 OCT4, REX1, UTE, TERT, SOX2, CX43 and CX45. They are
 negative for GATA2, GATA4, PDX1, NCAM, MSX1, FLT3, SSEA-1,
 TUBB3, NES, GFAP, and ROMEs. When confluent (18-10 days
 after plating), the ES cells from 4 X 6cm dishes were
 treated with 1 mg/ml collagenase, type IV
 (Invitrogen/GIBCO) for 5-10 min and gently scraped off
 with 5 ml pipette. RNA was purified with Trizol Reagent
 from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558
 (2001). [PMID: 11544199]) Double-stranded cDNAs were
 synthesized with an Oligo(dT) primer [Invitrogen:
 5'-pGACTAGTCTAGATCGAGCGCCCTTTT-3'] from
 3.4g of total RNA, treated with T4 DNA polymerase, and
 purified by ethanol-precipitation. The cDNAs were ligated
 to Lone-linker LL-Sal4, purified by phenol/chloroform
 extraction, and separated from free linkers by
 Centricon-100 column. Then, the cDNAs were amplified by
 long-range high fidelity PCR using Ex Taq polymerase

(Takara) with a primer Sal4-S for 25 cycles. The products were purified by phenol/chloroform extraction and Centricon-100 column. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The average insert size is about 3.6kb."

```

ORIGIN
Query Match      30.2%; Score 30.2; DB 6; Length 823;
Best Local Similarity 62.7%; Pred. No. 18;
Matches 47; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 22 GAGTCGCTAGTAGTCGCGAGCAAAATTTAAGCTACAACAGCAAGGCTTACCCAC 81
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 524 GAGCTAACTGAATTCGTATGCGAGCAGCATTTAAACATATTCCTAGTCAAGGACTGGATGGG 465
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 82 AATTGCATGAAGAAT 96
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 464 AAGTAAGTGAAGAAT 450
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 4
CC242469/c      1165 bp      DNA      linear      GSS 12-MAY-2003
LOCUS      CH261-11f8 RM1.1 CH261 Gallus gallus genomic clone CH261-11f8,
DEFINITION      genomic survey sequence.
ACCESSION      CC242469
VERSION      CC242469.1 GI:30569132
KEYWORDS      GSS.
SOURCE      Gallus gallus (chicken)
ORGANISM      Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE      1 (bases 1 to 1165)
AUTHORS      Wrenn,W., Graves,T., Mardis,E. and Wilson,R.
TITLE      Gallus gallus BAC End Reads
JOURNAL      Unpublished (2003)
COMMENT      Contact: Richard K. Wilson
              Genome Sequencing Center
              Washington University School of Medicine
              Email: submissions@watson.wustl.edu
              Insert Length: 182000 Std Error: 0.00
              Seq primer: RM1 TACGACTCACTATAGGAGGA
              Class: BAC ends
              High quality sequence start: 15
              High quality sequence stop: 738.
FEATURES
source      1..1165
              Location/Qualifiers
                /organism="Gallus gallus"
                /mol_type="genomic DNA"
                /strain="Red Jungle Fowl"
                /db_xref="taxon:9031"
                /clone="CH261-11f8"
                /sex="female"
                /cell_line="UCD001, inbred 256"
                /note="Vector: pTACBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
                CH261 Female Chicken library - For library and clone
                ordering information: http://www.choori.org/bacpac"

ORIGIN
Query Match      29.8%; Score 29.8; DB 8; Length 1165;
Best Local Similarity 58.4%; Pred. No. 26;
Matches 52; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Qy 7 CTTGCTTGTGTTGAGTCCCTGAGTAGTCGCGAGCAAAATTTAAGCTACAAAGG 66
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 676 CTTGCTTCTCTGTGGTGGAACTTATGTCCTCTGCTCCCTGCTGCAATTAATTTGCATAAAGG 617
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 67 CAGGCTTGACCGCAATTCGATGAAGAA 95
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 616 AAACACTTAAAGCCCAACTCCACAGAGAA 588
RESULT 5
CB387202      401 bp      mRNA      linear      EST 15-MAY-2003
LOCUS      OSTF076E6_1 AD-wrmcDNA Caenorhabditis elegans cDNA, mRNA sequence.
DEFINITION      CB387202
VERSION      CB387202.1 GI:30728912
KEYWORDS      EST.
SOURCE      Caenorhabditis elegans
ORGANISM      Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditoides; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE      1 (bases 1 to 401)
AUTHORS      Reboul,J., Vaglio,P., Rual,J.F., Lamesch,P., Martinez,M.,
              Armstrong,C.M., Li,S., Jacotot,L., Bertin,N., Janky,R., Moore,T.,
              Hudson,J.R., Hartley,J.L., Brasch,M.A., Vandenhaute,J., Boulton,S.,
              Endress,G.A., Jenna,S., Chevet,E., Papasotiropoulos,V.,
              Toliaas,P.P., Pracek,J., Snyder,M., Huang,R., Chance,M.R., Lee,H.,
              Doucette-Stamm,L., Hill,D.E. and Vidal,M.
              C. elegans ORFeome version 1.1: experimental verification of the
              genome annotation and resource for proteome-scale protein
              expression
              Nat. Genet. (2003) In press
              Contact: Vidal M
              Marc Vidal Laboratory
              Dana Farber Cancer Institute
              1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
              Tel: 617 632 5180
              Fax: 617 632 5739
              Email: Marc.Vidal@dfci.harvard.edu
              Sequence tag of Gateway entry clones. The primers used were
              designed on the predicted protein encoding ORF. C. elegans ORFeome
              cloning project : Contact david.hill@dfci.harvard.edu or
              marc.vidal@dfci.harvard.edu
              POLYA=No.
FEATURES
source      1..401
              Location/Qualifiers
                /organism="Caenorhabditis elegans"
                /mol_type="mRNA"
                /strain="N2"
                /db_xref="taxon:6239"
                /sex="Hermaphrodite and male"
                /tissue_type="whole animal"
                /dev_stage="mixed stage"
                /clone_lib="AD-wrmcDNA"
                /note="The AD-wrmcDNA library was generated with poly(A) +
                RNA isolated from both hermaphrodite and male N2 worms of
                all larval stages, embryos, adults and dauers and the
                subsequent generation of cDNAs by poly(A) priming. The
                cDNAs were cloned into pPC86"
ORIGIN
Query Match      29.4%; Score 29.4; DB 6; Length 401;
Best Local Similarity 60.8%; Pred. No. 30;
Matches 48; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

Qy 12 TTGTGTGTGAGTCCGTGAGTAGTCGCGAGCAAAATTTAAGCTACAAAGGCAAGG 71
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 288 TTTTCTCTTGATTAACAACAGTCATGAGAGAGCTTTAATTAATTAAGTACAGGATAAG 347
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 72 CTTGACCGCAATTCGATG 90
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 348 ATTTCTTGACCAATTCATG 366
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 6
BQ310441/c
LOCUS      BQ310441
DEFINITION      MR0-BT4502-220601-202-b02 Homo sapiens cDNA, mRNA sequence.
ACCESSION      BQ310441
VERSION      BQ310441.1 GI:20853032

```

```

KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 531)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.P.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bata,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR0&t2=MR0-BT4502-
220601-202-b02&t3=2001-06-22&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 48
High quality sequence stop: 530.
Location/Qualifiers
FEATURES
source
1..531
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="BT4502"
/note="Organ: breast; Vector: puc18; Site 1: SmaI; Site 2:
SmaI; A mini-library was made by cloning products derived
from ORESSES PCR (U.S. Letters Patent application No.
196,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
ORIGIN
Query Match 29.4%; Score 29.4; DB 5; Length 531;
Best Local Similarity 60.8%; Pred. No. 31;
Matches 48; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
Qy 22 GAGGTGCTGAGTAGTCGCGCAGCAAAATTTAAAGCTACAAACGCAAGGCTTGACCGAC 81
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 152 GAGCTAACTGAATGTATGGGAGCAGCATTAAACATATCTTAGTCAGGACGAGATGGG 93
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 82 AATTCATGAAGAATCTGC 100
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 92 AAGTAAGTGAAGATAGGC 74
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 7
AQ946479/c 754 bp DNA linear GSS 27-JAN-2000
LOCUS Sheared DNA-4906.TR Sheared DNA Trypanosoma brucei genomic clone
DEFINITION Sheared DNA-4906, genomic survey sequence.
ACCESSION AQ946479
VERSION AQ946479.1 GI:6769744
KEYWORDS GSS.
SOURCE Trypanosoma brucei
ORGANISM Trypanosoma
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
Trypanosoma.
REFERENCE 1 (bases 1 to 754)
AUTHORS El-Sayed,N., Zhao,S., Zhao,H., Gill,S., Suh,E., Malek,J., Fujii,C.,
Gerrard,C., Leech,V., de Jong,P., Ullu,E., Melville,S.,
Doneison,J., Fraser,C. and Adams,M.
TITLE Determination of clone end sequences from Trypanosoma brucei GUTat
10.1 sheared DNA library
JOURNAL Unpublished (1999)
COMMENT Other_GSSs: Sheared DNA-4906.TF
Contact: Najib M. El-Sayed
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: nelsayed@tigr.org
Clones are derived from the Trypanosoma brucei GUTat 10.1 sheared
DNA library constructed at TIGR. Clones will be available for
distribution through ATCC. Sheared DNA end sequences search page:
http://www.tigr.org/tdb/mdb/tbdb/.
Seq primer: M13-Reverse
Class: Shotgun.
Location/Qualifiers
FEATURES
source
1..754
/organism="Trypanosoma brucei"
/mol_type="genomic DNA"
/strain="TREU927/4 GUTat 10.1"
/db_xref="taxon:5691"
/clone="Sheared DNA-4906"
/clone_lib="Sheared DNA"
/note="Vector: pUC18; Site 1: SmaI; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from a cloned population of
Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically
sheared to give a tight size distribution (approx 2 kb).
The v + i method used for the library construction is
described in detail in Smith, H.O. and Venter, J.C.
(Making small insert libraries for whole genome shotgun
sequencing projects. In genome sequencing: A Practical
Approach, eds. M. Vaudin and B. Borell, Oxford University
Press, 1999)."
ORIGIN
Query Match 29.4%; Score 29.4; DB 8; Length 754;
Best Local Similarity 70.9%; Pred. No. 33;
Matches 39; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
Qy 45 CAAATTTAAGCTTACAAACGCAAGGCTTGACCGCAATTCATGAGAAATCTG 99
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 165 CACAAATTAAGCTTCAAAAGGCAAGGCTTGACCGCAATTCATGAGAAATCTG 111
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 8
CC888514 426 bp DNA linear GSS 31-JUL-2003
LOCUS SALK_151964.54.50.x Arabidopsis thaliana TDNA insertion lines
DEFINITION Arabidopsis thaliana genomic clone SALK_151964.54.50.x, genomic
survey sequence.
ACCESSION CC888514
VERSION CC888514.1 GI:33365229
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1 (bases 1 to 426)
AUTHORS Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,
Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
Shinn,P., Zimmerman,J. and Ecker,J.R.
TITLE A Sequence-Indexed Library of Insertion Mutations in the
Arabidopsis Genome
JOURNAL Unpublished (2001)
COMMENT Contact: Joseph R. Ecker

```

Salk Institute Genomic Analysis Laboratory (SIGAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: eckerg@alk.edu
This is single pass sequence recovered from the left border of
TDNA. This sequence lies within an annotated intron of AL1903960.
Class: TDNA tagged.

FEATURES

source

Location/Qualifiers

1..426
/organism="Arabidopsis thaliana"
/mol_type="Genomic DNA"
/ecotype="Col-0"
/db_xref="taxon:3702"
/clone="SALK151964.54.50.x"
/note="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna_protocols.html"

ORIGIN

Query Match 29.2%; Score 29.2; DB 9; Length 426;
Best Local Similarity 64.2%; Pred. No. 35;
Matches 43; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 30 TGAGTAGTGGCGGACGACAAATTTAGCTACACAGGCGAGCTTGACGACAAATGGAT 89

DB 76 TTATTAGTTGGTGTTCAGANTTTAGCATCATCAATCAAGACTTGACCTACAGAAACAT 17

QY 90 GAAGAAAT 96

DB 16 GAAAAAT 10

RESULT 9

CK086063/C
LOCUS RG11 C07 Cucumber leaf Cucumis sativus cDNA, mRNA linear EST 01-DEC-2003

ACCESSION CK086063

VERSION CK086063.1 GI:38571123

KEYWORDS EST.

SOURCE Cucumis sativus (cucumber)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Cucurbitales; Cucurbitaceae; Cucumis.

REFERENCE

AUTHORS Grunet,R. and McGrath,M.

TITLE Development of genomic tools for cucumber (Cucumis sativus L.)

JOURNAL Unpublished (2003)

COMMENT Contact: Rebecca Grunet

Rebecca Grunet

Michigan State University

Horticulture Department, Michigan State University, East Lansing,

MI 48824, USA

Tel: 517 353 0890

Fax: 517 355 5191 x431

Email: grunet@msu.edu

Plate: RG11 row: C column: 07.

Location/Qualifiers

1..657

/organism="Cucumis sativus"

/mol_type="mRNA"

/strain="Straight 8"

/db_xref="taxon:3659"

/sex="monoecious"

/clone_lib="Cucumber leaf"

/note="Vector: pAD-GAL4; Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Query Match 29.0%; Score 29; DB 7; Length 657;
Best Local Similarity 63.8%; Pred. No. 44;
Matches 44; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 31 GAGTAGTGGCGGACGACAAATTTAGCTACACAGGCGAGCTTGACGACAAATGCAATG 90

DB 267 GAGGAGTGCAAGAACCAACTGAAGCCAGAGTGAAGAGAGGCTTGCAAGGTCAAG 208

QY 91 AAGAATCTG 99

DB 207 AAAAATTTG 199

RESULT 10

CNS032VI/C

LOCUS CNS032VI/C

DEFINITION Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone

207F02 of library G from Tetraodon nigroviridis, genomic survey

sequence.

ACCESSION AL225351.1 GI:7884242

VERSION AL225351

KEYWORDS GSS; genome survey sequence.

SOURCE Tetraodon nigroviridis

ORGANISM Tetraodon nigroviridis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;

Tetraodontidae; Tetraodontidae; Tetraodon.

REFERENCE 1

AUTHORS Roest Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,

Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,

Saurin,W. and Weissenbach,J.

Estimate of human gene number provided by genome-wide analysis

using Tetraodon nigroviridis DNA sequence

Nat. Genet. 25 (2), 235-238 (2000)

20296633

REFERENCE 2

AUTHORS Roest Crolius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,

Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,

Saurin,W., Bernot,A. and Weissenbach,J.

Characterization and repeat analysis of the compact genome of the

freshwater pufferfish Tetraodon nigroviridis

Genome Res. 10 (7), 939-949 (2000)

20359837

PUBMED 10899143

REFERENCE 3 (bases 1 to 877)

Genoscope.

Direct Submission

Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : secref@genoscope.cns.fr)

- Web : www.genoscope.cns.fr

This sequence is a single read and was generated as part of a large

scale clone-end sequencing project of the Tetraodon nigroviridis

genome. For more information, please take a look at

<http://www.genoscope.cns.fr/Tetraodon>.

FEATURES

source

1..877

/organism="Tetraodon nigroviridis"

/mol_type="genomic DNA"

/db_xref="taxon:99883"

/clone="207F02"

/clone_lib="G"

/note="Genoscope sequence.ID : COAG207DC01SP1-end :

PUC-Ori"

ORIGIN

Query Match 29.0%; Score 29; DB 9; Length 877;

Best Local Similarity 63.8%; Pred. No. 47;

Matches 44; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 14 GTGTGTTGGAGTGGCTGAGTAGTGGCGGACGACAAATTTAGCTACACAGGCGAGGCT 73

```

Db 199 GGGTCCTGGGGGGCGCTCGCGTGCTCAAGCATTAACATCTACACATGGAAAGTTA 140
Qy 74 TGACCGGACA 82
Db 139 TGAAGAGAGA 131

RESULT 11
F32722/c
LOCUS
DEFINITION HSPD25699 HM3 Homo sapiens cDNA clone s3000037G06, mRNA sequence.
ACCESSION F32722
VERSION F32722.1 GI:4818348
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Lanfranchi,G., Muraro,T., Caldara,F., Pacchioni,B., Pallavicini,A.,
Pandofo,D., Toppo,S., Trevisan,S., Scarso,S. and Valle,G.
TITLE Identification of 4370 expressed sequence tags from a
3'-end-specific cDNA library of human skeletal muscle by DNA
sequencing and filter hybridization
JOURNAL Genome Res. 6 (1), 35-42 (1996)
MEDLINE 96276048
PUBMED 8691137
COMMENT Contact: Valle G.
CRIBI Biotechnology Centre
University of Padua
Via Trieste 75, 35121 Padua, Italy
ABI Chromatograms and other information are available on WWW at
http://group.bio.unipd.it.
FEATURES
source
1..340
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="s3000037G06"
/sex="female"
/tissue_type="pectoral muscle (after mastectomy)"
/clone_lib="HM3"
/note="Vector: pcDNAII (Invitrogen); Site 1: BstXI;
Site 2: NotI; The library was constructed by G.
Lanfranchi. This library is not subtracted nor normalized.
The first strand cDNA was primed with a biotinylated
oligo-dT-NotI primer
(5'-biotin-AACCGGCTCGAGCGCGCTTTT-3'). The
ds cDNA was sonicated and size-selected in the range
350-550 bp. The 3' specific fragments were selected by
streptavidin coated magnetic beads, ligated to
non-palindromic BstXI adapters, NotI digested and
directionally cloned into BstXI-NotI cut pcDNAII vector."
ORIGIN
Query Match 28.6%; Score 28.6; DB 7; Length 340;
Best Local Similarity 61.3%; Pred. No. 55;
Matches 46; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
Qy 22 GAGTCGCTGAGTAGTCGCGCAGCAAAATTTAAGCTACAACAGCGAGGTTGACCGAC 81
Db 125 GAGCTAACTGAATGTATGGAGCAGCATTTAACAATATCTCTAGTCAAGGACAGGATGGG 66
Qy 82 AATTGCATGAAGAAT 96
Db 65 AAGTAAGTGAAGAAT 51

Query Match 28.6%; Score 28.6; DB 7; Length 340;
Best Local Similarity 61.3%; Pred. No. 55;
Matches 46; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
Qy 22 GAGTCGCTGAGTAGTCGCGCAGCAAAATTTAAGCTACAACAGCGAGGTTGACCGAC 81
Db 125 GAGCTAACTGAATGTATGGAGCAGCATTTAACAATATCTCTAGTCAAGGACAGGATGGG 66
Qy 82 AATTGCATGAAGAAT 96
Db 65 AAGTAAGTGAAGAAT 51

RESULT 12
AA962465
LOCUS
DEFINITION 0091e05.s1 NCI_CGAP_Kid5 Homo sapiens cDNA clone IMAGE:1573568 3',

```

```

mRNA sequence.
AA962465
VERSION AA962465.1 GI:3134629
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 408)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-f@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
Insert Length: 1045 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham.
FEATURES
source
Location/Qualifiers
1..408
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1573568"
/tissue_type="2 pooled tumors (clear cell type)"
/clone_lib="NCI_CGAP_Kid5"
/lab_host="DHI08"
/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
AATGGAAGAAATCGCGCGCAATATTTT-3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo. "
ORIGIN
Query Match 28.6%; Score 28.6; DB 1; Length 408;
Best Local Similarity 61.3%; Pred. No. 57;
Matches 46; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
Qy 22 GAGTCGCTGAGTAGTCGCGCAGCAAAATTTAAGCTACAACAGCGAGGTTGACCGAC 81
Db 243 GAGCTAACTGAATGTATGGAGCAGCATTTAACAATATCTCTAGTCAAGGACAGGATGGG 302
Qy 82 AATTGCATGAAGAAT 96
Db 303 AAGTAAGTGAAGAAT 317

RESULT 13
CN386744/c
LOCUS
DEFINITION 328755673 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CN386744
VERSION CN386744.1 GI:47374339
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 436)
Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
Lebkowski,J and Stanton,L.W.

```

```

TITLE      Transcriptome characterization elucidates signaling networks that
JOURNAL    control human ES cell growth and differentiation
COMMENT    Nat. Biotechnol. 22 (6), 707-716 (2004)
           Contact: Brandenberger R
           Regenerative Medicine
           Genon Corporation
           230 Constitution Drive, Menlo Park, CA 94025, USA
           Tel: 650 473 8658
           Fax: 650 473 7760
           Email: rbrandenberger@genon.com
           Insert Length: 436 Std Error: 0.00.
FEATURES   Location/Qualifiers
source     1..436
           /organism="Homo sapiens"
           /mol_type="mRNA"
           /db_xref="taxon:9606"
           /tissue type="embryonic stem cells, embryoid bodies
           derived from H1, H7 and H9 cells"
           /clone_lib="GRN_EB"
           /note="Oligo dt primed, full-length enriched cDNA library
           from embryoid body outgrowths derived from HES cell lines
           H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
           conditions."

ORIGIN
Query Match      28.6%; Score 28.6; DB 7; Length 436;
Best Local Similarity 61.3%; Pred. No. 57;
Matches 46; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy 22 GAGTCGCTGAGTAGTCCGCGAGCAAAATTTAAAGCTACAACAGCAAGGCTTGACCGAC 81
Db 133 GAGCTAACTGAATTGTATGGGAGCAGCATTAAACATATTCCTAGTCAAGGACAGGATGGG 74

Qy 82 AATTGCATGAAGAT 96
Db 73 AAGTAAGTGAAGAT 59

RESULT 14
LOCUS      CB161201
DEFINITION K-EST0221011 L18POOL1n1 Homo sapiens cDNA clone L18POOL1n1-16-H03
5', mRNA sequence.
ACCESSION  CB161201
VERSION    CB161201.1 GI:28147327
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 514)
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
21C Frontier Korean EST Project 2001
JOURNAL    Unpublished (2002)
COMMENT    Contact: Kim YS
           Genome Research Center
           Korea Research Institute of Bioscience & Biotechnology
           52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
           Tel: +82-42-860-4470
           Fax: +82-42-860-4409
           Email: yongsung@mail.kribb.re.kr
           Plate: 16 row: H column: 03
           High quality sequence stop: 514.
           Location/Qualifiers
           1..514
           /organism="Homo sapiens"
           /mol_type="mRNA"
           /db_xref="taxon:9606"
           /clone="L18POOL1n1-16-H03"
           /cell_line="SNU-354+Cho-CK+Choi-CK+HLK-3"
           /lab_host="Top10P"

TITLE      Transcriptome characterization elucidates signaling networks that
JOURNAL    control human ES cell growth and differentiation
COMMENT    Nat. Biotechnol. 22 (6), 707-716 (2004)
           Contact: Brandenberger R
           Regenerative Medicine
           Genon Corporation
           230 Constitution Drive, Menlo Park, CA 94025, USA
           Tel: 650 473 8658
           Fax: 650 473 7760
           Email: rbrandenberger@genon.com
           Insert Length: 436 Std Error: 0.00.
FEATURES   Location/Qualifiers
source     1..436
           /organism="Homo sapiens"
           /mol_type="mRNA"
           /db_xref="taxon:9606"
           /clone="L18POOL1n1-16-H03"
           /cell_line="SNU-354+Cho-CK+Choi-CK+HLK-3"
           /lab_host="Top10P"

```

```

/clone_lib="L18POOL1n1"
/note="Organ: Liver; Vector: pT73-Pac; Site 1: EcoRI;
Site 2: NotI; The library was contributed by the Soares
laboratory and it was constructed as described by Bonaldo,
M.F., Lennon, G. and Soares, M.B. (1996), Genome Research
6(9): 791-806. RNA was prepared from harvested cell
culture."

ORIGIN
Query Match      28.6%; Score 28.6; DB 6; Length 514;
Best Local Similarity 61.3%; Pred. No. 59;
Matches 46; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy 22 GAGTCGCTGAGTAGTCCGCGAGCAAAATTTAAAGCTACAACAGCAAGGCTTGACCGAC 81
Db 261 GAGCTAACTGAATTGTATGGGAGCAGCATTAAACATATTCCTAGTCAAGGACAGGATGGG 320

Qy 82 AATTGCATGAAGAT 96
Db 321 AAGTAAGTGAAGAT 335

RESULT 15
LOCUS      CB161182
DEFINITION K-EST0220988 L18POOL1n1 Homo sapiens cDNA clone L18POOL1n1-16-F04
5', mRNA sequence.
ACCESSION  CB161182
VERSION    CB161182.1 GI:28147308
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 530)
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
21C Frontier Korean EST Project 2001
JOURNAL    Unpublished (2002)
COMMENT    Contact: Kim YS
           Genome Research Center
           Korea Research Institute of Bioscience & Biotechnology
           52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
           Tel: +82-42-860-4470
           Fax: +82-42-860-4409
           Email: yongsung@mail.kribb.re.kr
           Plate: 16 row: F column: 04
           High quality sequence stop: 530.
           Location/Qualifiers
           1..530
           /organism="Homo sapiens"
           /mol_type="mRNA"
           /db_xref="taxon:9606"
           /clone="L18POOL1n1-16-F04"
           /cell_line="SNU-354+Cho-CK+Choi-CK+HLK-3"
           /lab_host="Top10P"
           /clone_lib="L18POOL1n1"
           /note="Organ: Liver; Vector: pT73-Pac; Site 1: EcoRI;
           Site 2: NotI; The library was contributed by the Soares
           laboratory and it was constructed as described by Bonaldo,
           M.F., Lennon, G. and Soares, M.B. (1996), Genome Research
           6(9): 791-806. RNA was prepared from harvested cell
           culture."

FEATURES   Location/Qualifiers
source     1..530
           /organism="Homo sapiens"
           /mol_type="mRNA"
           /db_xref="taxon:9606"
           /clone="L18POOL1n1-16-F04"
           /cell_line="SNU-354+Cho-CK+Choi-CK+HLK-3"
           /lab_host="Top10P"

ORIGIN
Query Match      28.6%; Score 28.6; DB 6; Length 530;
Best Local Similarity 61.3%; Pred. No. 59;
Matches 46; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy 22 GAGTCGCTGAGTAGTCCGCGAGCAAAATTTAAAGCTACAACAGCAAGGCTTGACCGAC 81
Db 261 GAGCTAACTGAATTGTATGGGAGCAGCATTAAACATATTCCTAGTCAAGGACAGGATGGG 320

```

Qy 82 AATTGCATGAAGAAT 96
|||
Db 321 AAGTAAGTGAAGAAT 335
|||

Search completed: July 14, 2005, 23:23:14
Job time : 960.146 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:39:07 ; Search time 756.618 Seconds
(without alignments)
6468.225 Million cell updates/sec

Title: US-09-482-682-64_COPY_3565_3665
Perfect score: 101
Sequence: 1 gcgtgacgcgtacacttgc.....ttccccgcgaagctctaaat 101

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	101	100.0	299	6	CQ815667
C 2	101	100.0	460	6	I08196
C 3	101	100.0	461	6	I05488
C 4	101	100.0	461	6	I08820
C 5	101	100.0	461	6	I16794
C 6	101	100.0	470	6	A60961
C 7	101	100.0	470	6	A60977
C 8	101	100.0	470	6	AR369178
C 9	101	100.0	470	6	AR369188
C 10	101	100.0	470	6	AR476628
C 11	101	100.0	470	6	AR476638
C 12	101	100.0	470	6	AR487254
C 13	101	100.0	470	6	AR487264
C 14	101	100.0	472	6	AR160383
C 15	101	100.0	472	6	BD194797
C 16	101	100.0	472	6	AX482611
C 17	101	100.0	698	6	A85395
C 18	101	100.0	698	6	AR154888
C 19	101	100.0	698	6	E65413

C 20	101	100.0	729	6	BD237245
C 21	101	100.0	729	6	AR240810
C 22	101	100.0	730	6	BD237246
C 23	101	100.0	730	6	AR240811
C 24	101	100.0	757	6	BD221144
C 25	101	100.0	801	6	AR370661
C 26	101	100.0	825	6	AX284008
C 27	101	100.0	858	6	AX752973
C 28	101	100.0	902	6	AR505921
C 29	101	100.0	910	9	LDO244004
C 30	101	100.0	1217	6	AR036903
C 31	101	100.0	1217	6	AR141142
C 32	101	100.0	1217	6	AR181917
C 33	101	100.0	1218	5	AJ719748
C 34	101	100.0	1695	12	AB003139
C 35	101	100.0	2000	6	I01987
C 36	101	100.0	2029	12	AY733067
C 37	101	100.0	2175	12	AY733070
C 38	101	100.0	2322	6	AX670965
C 39	101	100.0	2326	12	AF143508
C 40	101	100.0	2380	6	A60985
C 41	101	100.0	2380	6	AR369195
C 42	101	100.0	2380	6	AR476645
C 43	101	100.0	2380	6	AR487271
C 44	101	100.0	2383	12	CVU50331
C 45	101	100.0	2402	6	AX670963

ALIGNMENTS

RESULT 1
LOCUS CQ815667/c 299 bp DNA linear PAT 03-JUN-2004
DEFINITION Sequence 65 from Patent WO2004042036.
ACCESSION CQ815667
VERSION CQ815667.1 GI:48144221
KEYWORDS
SOURCE Zygosaccharomyces bailii
ORGANISM Zygosaccharomyces bailii
Eukaryota; Fungi; Ascomycota; Saccharomycetaceae; Zygosaccharomycetes;
Saccharomycetales; Saccharomycetaceae; Zygosaccharomycetes.
REFERENCE 1
AUTHORS Porro,D., Branduardi,P., Valli,M. and Alberghina,L.
TITLE Process for expression and secretion of proteins by the non-conventional yeast zygosaccharomyces bailii
JOURNAL Patent: WO 2004042036-A 65 21-MAY-2004;
FEATURES
source Porro, Danilo (IT)
Location/Qualifiers
1..299
/organism="Zygosaccharomyces bailii"
/mol_type="unassigned DNA"
/db_xref="taxon:4954"

ORIGIN
Query Match: 100.0%; Score 101; DB 6; Length 299;
Best Local Similarity 100.0%; Pred. No. 6.6e-16;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTTCCTTCGCTTCTTCCTTCCT 60
Db 156 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTTCCTTCGCTTCTTCCTTCCT 97
Qy 61 TTCTGCGCAGTTCGCGCGCTTTCCTGCTCAAGCTCTAAAT 101
Db 96 TTCTGCGCAGTTCGCGCGCTTTCCTGCTCAAGCTCTAAAT 56

RESULT 2
LOCUS I08196/c 460 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 1 from Patent EP 0356130.
ACCESSION I08196

```
VERSION I08196.1 GI:589091
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 460)
AUTHORS Weber,S.C., Holzschu,D.Lc. and Lalik,P.Hc.
TITLE A mobile fl phage single-strand DNA origin of replication
JOURNAL Patent: EP 0356130-A2 1 28-FEB-1990;
FEATURES Location/Qualifiers
          1..460
          /organism="unknown"
          /mol_type="unassigned DNA"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 460;
Best Local Similarity 100.0%; Pred. No. 6.2e-16;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTTTCGCTTCTTCCCTTCCCT 60
Db 406 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTTTCGCTTCTTCCCTTCCCT 347
Qy 61 TTCTCGCACGTTCCGCGGCTTTCGCGTCAAGCTCTAAAT 101
Db 346 TTCTCGCACGTTCCGCGGCTTTCGCGTCAAGCTCTAAAT 306
RESULT 3
I05488
LOCUS I05488 461 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 1 from Patent EP 0286200.
ACCESSION I05488
VERSION I05488.1 GI:590716
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 461)
AUTHORS Sorge,J.A.M., Huse,W.M. and Short,J.M.
TITLE DNA cloning vector with in vivo excisable plasmids
JOURNAL Patent: EP 0286200-A2 1 12-OCT-1988;
FEATURES Location/Qualifiers
          1..461
          /organism="unknown"
          /mol_type="unassigned DNA"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 461;
Best Local Similarity 100.0%; Pred. No. 6.2e-16;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTTTCGCTTCTTCCCTTCCCT 60
Db 54 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTTTCGCTTCTTCCCTTCCCT 113
Qy 61 TTCTCGCACGTTCCGCGGCTTTCGCGTCAAGCTCTAAAT 101
Db 114 TTCTCGCACGTTCCGCGGCTTTCGCGTCAAGCTCTAAAT 154
RESULT 4
I08820
LOCUS I08820 461 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 1 from Patent WO 8805085.
ACCESSION I08820
VERSION I08820.1 GI:588470
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 461)
AUTHORS Huse,W., Sorge,J.A. and Short,J.M.
JOURNAL Patent: WO 8805085-A 1 14-JUL-1988;
FEATURES Location/Qualifiers
          1..461
          /organism="unknown"
          /mol_type="unassigned DNA"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 461;
Best Local Similarity 100.0%; Pred. No. 6.2e-16;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTTTCGCTTCTTCCCTTCCCT 60
Db 54 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTTTCGCTTCTTCCCTTCCCT 113
Qy 61 TTCTCGCACGTTCCGCGGCTTTCGCGTCAAGCTCTAAAT 101
Db 114 TTCTCGCACGTTCCGCGGCTTTCGCGTCAAGCTCTAAAT 154
RESULT 5
I16794
LOCUS I16794 461 bp DNA linear PAT 03-APR-1996
DEFINITION Sequence 2 from patent US 5478731.
ACCESSION I16794
VERSION I16794.1 GI:1251702
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 461)
AUTHORS Short,J.M.
TITLE Polycos vectors
JOURNAL Patent: US 5478731-A 2 26-DEC-1995;
FEATURES Location/Qualifiers
          1..461
          /organism="unknown"
          /mol_type="unassigned DNA"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 461;
Best Local Similarity 100.0%; Pred. No. 6.2e-16;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTTTCGCTTCTTCCCTTCCCT 60
Db 54 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTTTCGCTTCTTCCCTTCCCT 113
Qy 61 TTCTCGCACGTTCCGCGGCTTTCGCGTCAAGCTCTAAAT 101
Db 114 TTCTCGCACGTTCCGCGGCTTTCGCGTCAAGCTCTAAAT 154
RESULT 6
A60961
LOCUS A60961 470 bp DNA linear PAT 06-MAR-1998
DEFINITION Sequence 270 from Patent WO9708320.
ACCESSION A60961
VERSION A60961.1 GI:3715496
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1
AUTHORS Knapik,A., Pack,P., Ilag,V., Ge,L., Moroney,S. and Plueckthun,A.
TITLE PROTEIN/(POLY)PEPTIDE LIBRARIES
JOURNAL Patent: WO 9708320-A 270 06-MAR-1997;
FEATURES Location/Qualifiers
          1..470
          /organism="unidentified"
          /mol_type="unassigned DNA"
          /db_xref="taxon:32644"
```

ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 470;
Best Local Similarity 100.0%; Pred. No. 6.1e-16;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTTTCGCTTCTTCCCTTCCCT 60
Db 58 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTTTCGCTTCTTCCCTTCCCT 117

Qy 61 TTCTCGCCACGTTCCGCGGCTTCCCGCTCAAGCTCTAAAT 101
Db 118 TTCTCGCCACGTTCCGCGGCTTCCCGCTCAAGCTCTAAAT 158

RESULT 7

LOCUS A60977 470 bp DNA linear PAT 06-MAR-1998
DEFINITION Sequence 286 from Patent WO9708320.
ACCESSION A60977
VERSION A60977.1 GI:3715509

KEYWORDS unidentified

SOURCE unidentified

ORGANISM unclassified.

REFERENCE 1
AUTHORS Knappik, A., Pack, P., Ilag, V., Ge, L., Moroney, S. and Plueckthun, A.

TITLE PROTEIN/(POLY)PEPTIDE LIBRARIES

JOURNAL PATENT: WO 9708320-A 286 06-MAR-1997;

MORPHOSYS PROTEINOPTIMIERUNG (DE)

FEATURES

source 1
Location/Qualifiers
1..470
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 470;
Best Local Similarity 100.0%; Pred. No. 6.1e-16;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTTTCGCTTCTTCCCTTCCCT 60
Db 58 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTTTCGCTTCTTCCCTTCCCT 117

Qy 61 TTCTCGCCACGTTCCGCGGCTTCCCGCTCAAGCTCTAAAT 101
Db 118 TTCTCGCCACGTTCCGCGGCTTCCCGCTCAAGCTCTAAAT 158

RESULT 8

LOCUS AR369178 470 bp DNA linear PAT 12-SEP-2003
DEFINITION Sequence 270 from patent US 6300064.
ACCESSION AR369178

VERSION AR369178.1 GI:34605134

KEYWORDS Unknown.

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 470)

AUTHORS Knappik, A., Pack, P., Ge, L., Moroney, S. and Plueckthun, A.

TITLE Protein/(poly)peptide libraries

JOURNAL PATENT: US 6300064-A 270 09-OCT-2001;

Location/Qualifiers

FEATURES

source 1
Location/Qualifiers
1..470
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 470;
Best Local Similarity 100.0%; Pred. No. 6.1e-16;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy

Db 58 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTTTCGCTTCTTCCCTTCCCT 117

Qy 61 TTCTCGCCACGTTCCGCGGCTTCCCGCTCAAGCTCTAAAT 101
Db 118 TTCTCGCCACGTTCCGCGGCTTCCCGCTCAAGCTCTAAAT 158

RESULT 9

LOCUS AR369188 470 bp DNA linear PAT 12-SEP-2003
DEFINITION Sequence 286 from patent US 6300064.
ACCESSION AR369188

VERSION AR369188.1 GI:34605144

KEYWORDS Unknown.

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 470)

AUTHORS Knappik, A., Pack, P., Ge, L., Moroney, S. and Plueckthun, A.

TITLE Protein/(poly)peptide libraries

JOURNAL PATENT: US 6300064-A 286 09-OCT-2001;

Location/Qualifiers

FEATURES

source 1
Location/Qualifiers
1..470
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 470;
Best Local Similarity 100.0%; Pred. No. 6.1e-16;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTTTCGCTTCTTCCCTTCCCT 60
Db 58 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTTTCGCTTCTTCCCTTCCCT 117

Qy 61 TTCTCGCCACGTTCCGCGGCTTCCCGCTCAAGCTCTAAAT 101
Db 118 TTCTCGCCACGTTCCGCGGCTTCCCGCTCAAGCTCTAAAT 158

RESULT 10

LOCUS AR476628 470 bp DNA linear PAT 14-MAY-2004
DEFINITION Sequence 270 from patent US 6696248.
ACCESSION AR476628

VERSION AR476628.1 GI:47233721

KEYWORDS Unknown.

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 470)

AUTHORS Knappik, A., Pack, P., Ge, L., Moroney, S. and Plueckthun, A.

TITLE Protein/(poly)peptide libraries

JOURNAL PATENT: US 6696248-A 270 24-FEB-2004;

Location/Qualifiers

FEATURES

source 1
Location/Qualifiers
1..470
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 101; DB 6; Length 470;
Best Local Similarity 100.0%; Pred. No. 6.1e-16;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTTTCGCTTCTTCCCTTCCCT 60
Db 58 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTTTCGCTTCTTCCCTTCCCT 117

Qy 61 TTCTCGCCACGTTCCGCGGCTTCCCGCTCAAGCTCTAAAT 101
Db 118 TTCTCGCCACGTTCCGCGGCTTCCCGCTCAAGCTCTAAAT 158

Db 118 TTCTCGCCACGTTCCGCGGCTTTCCCGCTCAAGCTCTAAAT 158

RESULT 11
AR476638 LOCUS 470 bp DNA linear PAT 14-MAY-2004
DEFINITION Sequence 286 from patent US 6696248.
ACCESSION AR476638
VERSION AR476638.1 GI:47233731
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 470)
AUTHORS Knappik,A., Pack,P., Ge,L., Moroney,S. and Pluckthun,A.
TITLE Protein/(poly)peptide libraries
JOURNAL Patent: US 6696248-A 286 24-FEB-2004;
FEATURES Location/Qualifiers
source 1..470
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 470;
Best Local Similarity 100.0%; Pred. No. 6.1e-16;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCTTTCGCTTTCTCCCTTCCT 60
Db 58 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCTTTCGCTTTCTCCCTTCCT 117

Qy 61 TTCTCGCCACGTTCCGCGGCTTTCCCGTCAAGCTCTAAAT 101
Db 118 TTCTCGCCACGTTCCGCGGCTTTCCCGTCAAGCTCTAAAT 158

RESULT 12
AR487254 LOCUS 470 bp DNA linear PAT 14-MAY-2004
DEFINITION Sequence 270 from patent US 6706484.
ACCESSION AR487254
VERSION AR487254.1 GI:47252205
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 470)
AUTHORS Knappik,A., Pack,P., Ge,L., Moroney,S. and Pluckthun,A.
TITLE Protein/(poly)peptide libraries
JOURNAL Patent: US 6706484-A 270 16-MAR-2004;
FEATURES Location/Qualifiers
source 1..470
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 470;
Best Local Similarity 100.0%; Pred. No. 6.1e-16;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCTTTCGCTTTCTCCCTTCCT 60
Db 58 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCTTTCGCTTTCTCCCTTCCT 117

Qy 61 TTCTCGCCACGTTCCGCGGCTTTCCCGTCAAGCTCTAAAT 101
Db 118 TTCTCGCCACGTTCCGCGGCTTTCCCGTCAAGCTCTAAAT 158

RESULT 13
AR487264 LOCUS 470 bp DNA linear PAT 14-MAY-2004
DEFINITION Sequence 286 from patent US 6706484.

ACCESSION AR487264 GI:47252215
VERSION AR487264.1
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 470)
AUTHORS Knappik,A., Pack,P., Ge,L., Moroney,S. and Pluckthun,A.
TITLE Protein/(poly)peptide libraries
JOURNAL Patent: US 6706484-A 286 16-MAR-2004;
FEATURES Location/Qualifiers
source 1..470
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 470;
Best Local Similarity 100.0%; Pred. No. 6.1e-16;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCTTTCGCTTTCTCCCTTCCT 60
Db 58 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCTTTCGCTTTCTCCCTTCCT 117

Qy 61 TTCTCGCCACGTTCCGCGGCTTTCCCGTCAAGCTCTAAAT 101
Db 118 TTCTCGCCACGTTCCGCGGCTTTCCCGTCAAGCTCTAAAT 158

RESULT 14
AR160383 LOCUS 472 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 5 from patent US 6255071.
ACCESSION AR160383
VERSION AR160383.1 GI:16224205
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 472)
AUTHORS Beach,D.H., Hannon,G.J., Conklin,D. and Sun,P.
TITLE Mammalian viral vectors and their uses
JOURNAL Patent: US 6255071-A 5 03-JUL-2001;
FEATURES Location/Qualifiers
source 1..472
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 472;
Best Local Similarity 100.0%; Pred. No. 6.1e-16;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCTTTCGCTTTCTCCCTTCCT 60
Db 63 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCTTTCGCTTTCTCCCTTCCT 122

Qy 61 TTCTCGCCACGTTCCGCGGCTTTCCCGTCAAGCTCTAAAT 101
Db 123 TTCTCGCCACGTTCCGCGGCTTTCCCGTCAAGCTCTAAAT 163

RESULT 15
BD194797 LOCUS 472 bp DNA linear PAT 17-JUL-2003
DEFINITION Viral vectors and their uses.
ACCESSION BD194797
KEYWORDS BD194797.1 GI:33004545
SOURCE JP 2002514054-A/6.
ORGANISM unidentified
REFERENCE 1 (bases 1 to 472)

Search completed: July 14, 2005, 14:03:36
Job time : 758.618 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:35:42 ; Search time 142.398 Seconds
(without alignments)
4198.742 Million cell updates/sec

Title: US-09-482-682-64_COPY_3565_3665

Perfect score: 101

Sequence: 1 gcgtgaccgtacactgccc.....ttcccgctgaagctcttaaat 101

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_16Dec04.*

1: Geneseq1980s.*

2: Geneseq1990s.*

3: Geneseq2000s.*

4: Geneseq2001as.*

5: Geneseq2001bs.*

6: Geneseq2002as.*

7: Geneseq2002bs.*

8: Geneseq2003as.*

9: Geneseq2003bs.*

10: Geneseq2003cs.*

11: Geneseq2003ds.*

12: Geneseq2004as.*

13: Geneseq2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101	100.0	288	4	AAS333391
2	101	100.0	288	4	AAS333392
3	101	100.0	288	4	AAS333392
4	101	100.0	288	4	AAS333392
5	101	100.0	288	4	AAS333391
6	101	100.0	288	4	AAS333391
7	101	100.0	288	4	AAS333391
8	101	100.0	288	4	AAS333391
9	101	100.0	288	4	AAS333391
10	101	100.0	288	4	AAS333391
11	101	100.0	288	4	AAS333391
12	101	100.0	288	4	AAS333391
13	101	100.0	288	4	AAS333391
14	101	100.0	288	4	AAS333391
15	101	100.0	288	4	AAS333391
16	101	100.0	288	4	AAS333391
17	101	100.0	288	4	AAS333391
18	101	100.0	288	4	AAS333391
19	101	100.0	288	4	AAS333391
20	101	100.0	288	4	AAS333391

21	101	100.0	359	8	ABX38602	Abx38602 Bovine ES
22	101	100.0	359	8	ABX41939	Abx41939 Bovine ES
23	101	100.0	359	8	ABX42410	Abx42410 Bovine ES
24	101	100.0	361	8	ABX46893	Abx46893 Bovine ES
25	101	100.0	361	8	ABX48545	Abx48545 Bovine ES
26	101	100.0	362	8	ABX46167	Abx46167 Bovine ES
27	101	100.0	362	8	ABX38860	Abx38860 Bovine ES
28	101	100.0	364	8	ABX46417	Abx46417 Bovine ES
29	101	100.0	365	8	ABX39110	Abx39110 Bovine ES
30	101	100.0	365	8	ABX41262	Abx41262 Bovine ES
31	101	100.0	365	8	ABX46903	Abx46903 Bovine ES
32	101	100.0	365	8	ABX38856	Abx38856 Bovine ES
33	101	100.0	366	8	ABX39785	Abx39785 Bovine ES
34	101	100.0	366	8	ABX43577	Abx43577 Bovine ES
35	101	100.0	366	8	ABX43580	Abx43580 Bovine ES
36	101	100.0	366	8	ABX45460	Abx45460 Bovine ES
37	101	100.0	366	8	ABX44973	Abx44973 Bovine ES
38	101	100.0	366	8	ABX48549	Abx48549 Bovine ES
39	101	100.0	366	8	ABX35989	Abx35989 Bovine ES
40	101	100.0	367	8	ABX37877	Abx37877 Bovine ES
41	101	100.0	367	8	ABX39577	Abx39577 Bovine ES
42	101	100.0	368	8	ABX42411	Abx42411 Bovine ES
43	101	100.0	368	8	ABX35306	Abx35306 Bovine ES
44	101	100.0	369	8	ABX36699	Abx36699 Bovine ES
45	101	100.0	369	8	ABX40761	Abx40761 Bovine ES

ALIGNMENTS

RESULT 1

AAS333391
ID AAS333391 standard; DNA; 288 BP.

AC AAS333391;

XX 04-DEC-2001 (first entry)

DE DNA encoding human secreted protein, Seq ID NO 674.

XX Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;
XX rheumatoid arthritis; antiarteriosclerotic; cardiant; vascular;
XX cerebroprotective; thrombolytic; antimicrobial; ophthalmological;
XX cystostatic; Alzheimer's disease; Parkinson's disease; human; cancer;
XX multiple sclerosis; cancer; hyperproliferative disorder; infection;
XX Gaucher's disease; neurological disease; cerebrovascular disorder;
XX thrombosis; wound healing; ds.

OS Homo sapiens.

PN WO200155326-A2.

XX 02-AUG-2001.

PD 17-JAN-2001; 2001WO-US001347.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

CC in samples (e.g. by enzyme linked immunosorbant assay (ELISA)). The
CC disorders include for example: immune/autoimmune diseases (e.g. HIV
CC (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis
CC and multiple sclerosis), cancers and hyperproliferative disorders (e.g.
CC melanomas, neoplasms of the breast or liver, Sezary syndrome and
CC Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,
CC Parkinson's disease and Charcot-Marie-Tooth disease), cardio-/
CC cerebrovascular disorders (e.g. cardiac arrest, tachycardia, angina and
CC thrombosis), infections caused by bacteria, viruses and fungi and ocular
CC disorders (e.g. corneal infections). (I) and (II), agonists, antagonists
CC and antibodies can also be used to promote wound healing, maintain organs
CC before transplantation, and support cell culture of primary tissues. PCR
CC AAS33043-AAS33486 represent human secreted protein coding sequences. PCR

Query Match 100.0%; Score 101; DB 4; Length 288;
Best Local Similarity 100.0%; Pred. No. 4.4e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CGGTGACCGCTACACTTCGCGCGCCCTAGCGCGCTTCCTTCCTTCCTTCCTTCCT 60
Db 33 CGGTGACCGCTACACTTCGCGCGCCCTAGCGCGCTTCCTTCCTTCCTTCCTTCCT 92
Qy 61 TTCTGCGCCACGTTCCCGCGCTTCCCGCTCAAGCTCTAAAT 101
Db 93 TTCTGCGCCACGTTCCCGCGCTTCCCGCTCAAGCTCTAAAT 133

RESULT 2

ID AAS33392 standard; DNA; 288 BP.

XX AC AAS333392;

XX DT 04-DEC-2001 (first entry)

XX DE DNA encoding human secreted protein, Seq ID No 675.

XX Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;
KW rheumatoid arthritis; antiarteriosclerotic; cardiac; vascular;
KW cerebroprotective; thrombolytic; antimicrobial; ophthalmological;
KW cystostatic; Alzheimer's disease; Parkinson's disease; cancer;
KW multiple sclerosis; cancer; hyperproliferative disorder; infection;
KW Gaucher's disease; neurological disease; cerebrovascular disorder;
KW thrombosis; wound healing; ds.

XX OS Homo sapiens.

XX PN WO200155326-A2.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001347.

XX 31-JAN-2000; 2000US-0179065P.

XX 04-FEB-2000; 2000US-0180628P.

XX 24-FEB-2000; 2000US-0184664P.

XX 02-MAR-2000; 2000US-0186350P.

XX 16-MAR-2000; 2000US-0189874P.

XX 17-MAR-2000; 2000US-0190076P.

XX 18-APR-2000; 2000US-0198123P.

XX 19-MAY-2000; 2000US-0205515P.

XX 07-JUN-2000; 2000US-0209467P.

XX 28-JUN-2000; 2000US-0214886P.

XX 30-JUN-2000; 2000US-0215135P.

XX 07-JUL-2000; 2000US-0216647P.

XX 07-JUL-2000; 2000US-0216880P.

PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225477P.
PR 14-AUG-2000; 2000US-0225575P.
PR 14-AUG-2000; 2000US-0225575P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.

PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249255P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-451931/48.
XX
XX New nucleic acids and polypeptides, useful for diagnosing, preventing or
XX treating medical conditions.
XX
XX Disclosure; SEQ ID NO 675; 753pp; English.
XX
XX The invention relates to novel isolated nucleic acid molecules (I)
XX encoding human secreted proteins (II). (I) and (II) are used to prevent,
XX treat or ameliorate a medical condition in e.g. humans, mice, rabbits,
XX goats, horses, cats, dogs, chickens or sheep. (I) and (II) may be used in
XX the prevention, treatment and diagnosis of diseases associated with
XX inappropriate expression of secreted proteins. (I) and complementary
XX sequences may also be used as DNA probes in diagnostic assays (e.g.
XX polymerase chain reactions (PCR)) to detect and quantitate the presence
XX of similar nucleic acid sequences in samples, and so which patients may
XX be in need of restorative therapy. (II) may also be used as antigens in
XX the production of antibodies and in assays to identify modulators
XX (agonists and antagonists) of the expression and activity of the secreted
XX proteins. The anti-(II) antibodies and antagonists may also be used to
XX down regulate expression and activity of (II). The anti-(II) antibodies
XX may also be used as diagnostic agents for detecting the presence of (II)
XX in samples (e.g. by enzyme linked immunosorbent assay (ELISA)). The
XX disorders include for example: immune/autoimmune diseases (e.g. HIV
XX (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis

CC and multiple sclerosis), cancers and hyperproliferative disorders (e.g.
CC melanomas, neoplasms of the breast or liver, Sezary syndrome and
CC Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,
CC Parkinson's disease and Charcot-Marie-Tooth disease), cardio-/
CC cerebrovascular disorders (e.g. cardiac arrest, tachycardia, angina and
CC thrombosis), infections caused by bacteria, viruses and fungi and ocular
CC disorders (e.g. corneal infections). (I) and (II), agonists, antagonists
CC and antibodies can also be used to promote wound healing, maintain organs
CC before transplantation, and support cell culture of primary tissues.
CC AAS33043-AAS33486 represent human secreted protein coding sequences, PCR

Query Match 100.0%; Score 101; DB 4; Length 288;

Best Local Similarity 100.0%; Pred. No. 4.4e-21;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGTGACCGGTACACTTGCAGCGCCCTAGCGCCGCTCTTCGTTTCTTCCCTTCCT 60
Db 33 GCGTGACCGGTACACTTGCAGCGCCCTAGCGCCGCTCTTCGTTTCTTCCCTTCCT 92
QY 61 TTCTCGCCAGTTTCGGCGGCTTTCCCGTCAAGCTCTAAAT 101
Db 93 TTCTCGCCAGTTTCGGCGGCTTTCCCGTCAAGCTCTAAAT 133

RESULT 3

AAK87482
ID AAK87482 standard; DNA; 288 BP.

XX AAK87482;

DT 07-NOV-2001 (first entry)

XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:42294.

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX cytostatic; gene therapy; vaccine; metastasis; ds.

XX Homo sapiens.

XX WO200157182-A2.

XX 09-AUG-2001.

XX 17-JAN-2001; 2001WO-US001354.

XX 31-JAN-2000; 2000US-0179065P.

XX 04-FEB-2000; 2000US-0180628P.

XX 24-FEB-2000; 2000US-0184664P.

XX 02-MAR-2000; 2000US-0186350P.

XX 16-MAR-2000; 2000US-0189874P.

XX 17-MAR-2000; 2000US-0190076P.

XX 18-APR-2000; 2000US-0198123P.

XX 19-MAY-2000; 2000US-0205515P.

XX 07-JUN-2000; 2000US-0209467P.

XX 28-JUN-2000; 2000US-0214886P.

XX 30-JUN-2000; 2000US-0215135P.

XX 07-JUL-2000; 2000US-0216647P.

XX 07-JUL-2000; 2000US-0216880P.

XX 11-JUL-2000; 2000US-0217487P.

XX 11-JUL-2000; 2000US-0217496P.

XX 14-JUL-2000; 2000US-0218290P.

XX 26-JUL-2000; 2000US-0220963P.

XX 26-JUL-2000; 2000US-0220964P.

XX 14-AUG-2000; 2000US-0224518P.

XX 14-AUG-2000; 2000US-0224519P.

XX 14-AUG-2000; 2000US-0225213P.

XX 14-AUG-2000; 2000US-0225214P.

XX 14-AUG-2000; 2000US-0225266P.

XX 14-AUG-2000; 2000US-0225267P.

XX 14-AUG-2000; 2000US-0225268P.

XX 14-AUG-2000; 2000US-0225270P.

XX 14-AUG-2000; 2000US-0225447P.

XX 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.

PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rojen CA, Barash SC, Ruben SM;

XX WPI; 2001-483426/52.

XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and metastasis.

XX Disclosure; SEQ ID NO 42294; 3071pp + Sequence Listing; English.

XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
XX amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patients own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting the
XX nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/haematopoietic-related diseases, especially
XX cancers and cancer metastases of haematopoietic-derived cells. AAK64703
XX to AAK87694 represent human immune/haematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAM82169
XX represent sequences used in the exemplification of the present invention

SQ Sequence 288 BP; 43 A; 87 C; 73 G; 85 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 4; Length 288;

Best Local Similarity 100.0%; Pred. No. 4.4e-21;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCCTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCTTCCTTCCTTCCTTCCT 60
|||||

Db 33 GCGTCACCGCTACACTTCCAGCGCCCTAGCGCCGCTCTTCGCTTTCTTCCTTCCT 92
Qy 61 TTCTCGCACGCTTCCGCGCTTTCCCGCTCAAGCTCTAAAT 101
Db 93 TTCTCGCACGCTTCCGCGCTTTCCCGCTCAAGCTCTAAAT 133

RESULT 4
AAK87481
ID AAK87481 standard; DNA; 288 BP.
XX AC AAK87481;
XX DT 07-NOV-2001 (first entry)
XX DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:42293.
XX KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX OS Homo sapiens.
XX FN WO200157182-A2.
XX PD 09-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US001354.
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 17-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 18-AUG-2000; 2000US-0225759P.
PR 22-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 05-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 06-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 12-SEP-2000; 2000US-0232081P.
PR 14-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-024617P.
PR 08-NOV-2000; 2000US-024647P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 17-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.

PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-483426/52.
DR
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
XX Disclosure; SEQ ID NO 42293; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAK62170 to AAK91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patient's own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention
XX
XX Sequence 288 BP; 43 A; 87 C; 73 G; 85 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 101; DB 4; Length 288;
Best Local Similarity 100.0%; Pred. No. 4.4e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCGTGACCGCTACACTTCCAGCGCCCTAGCGCGCTTCCTTCGCTTCTTCCTTCCT 60
Db 33 GCGTGACCGCTACACTTCCAGCGCCCTAGCGCGCTTCCTTCGCTTCTTCCTTCCT 92
QY 61 TTCTCGCACGTTCCCGGCTTTCCTTCGCTTCTTCCTTCCTTCCTTCCTTCCTTCCT 101
Db 93 TTCTCGCACGTTCCCGGCTTTCCTTCGCTTCTTCCTTCCTTCCTTCCTTCCTTCCT 133
RESULT 5
AAL07028
ID AAL07028 standard; DNA; 288 BP.
XX
XX AAL07028;
XX
XX 21-NOV-2001 (first entry)
XX
XX Human reproductive system related antigen DNA SEQ ID NO: 9716.
DE
XX

KW Human; reproductive system related antigen; reproductive system disorder;
KW cancer; gene therapy; ds.
XX Homo sapiens.
OS WO200155320-A2.
PN 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001339.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225577P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.

```
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-02355834P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-02355834P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-465570/50.
XX Isolated nucleic acid molecule encoding a reproductive system antigen is
XX used in preventing, treating or ameliorating a medical condition.
XX Disclosure; SEQ ID NO 9716; 1297pp + Sequence Listing; English.
XX The present invention provides the protein and coding sequences of a
XX number of human reproductive system related antigens. These can be used
XX in the prevention and treatment of reproductive system disorders,
XX CC including cancer. The present sequence is a genomic sequence encoding a
XX protein of the invention
XX SQ Sequence 288 BP; 43 A; 87 C; 73 G; 85 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 4; Length 288;
Best Local Similarity 100.0%; Pred. No. 4.4e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTTCCCGCTCAAGCTCTAAAT 101
Db 33 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTTCCCGCTCTAAAT 133
Qy 61 TTCTCGCCAGTTTCGGCGGCTTTCCTCCGCTCAAGCTCTAAAT 101
Db 93 TTCTCGCCAGTTTCGGCGGCTTTCCTCCGCTCAAGCTCTAAAT 133
RESULT 6
AAL07029
ID AAL07029 standard; DNA; 288 BP.
XX AAL07029;
XX 21-NOV-2001 (first entry)
XX Human reproductive system related antigen DNA SEQ ID NO: 9717.
XX Human; reproductive system related antigen; reproductive system disorder;
XX cancer; gene therapy; ds.
XX Homo sapiens.
XX WO200155320-A2.
XX 02-AUG-2001.
XX 17-JAN-2001; 2001WO-US001339.
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180828P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 11-JUL-2000; 2000US-0217487P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
```


QY 61 TTCTCGCCAGCTTCGCGGCTTTCCCGTCAAGCTCTAAAT 101
 |||||
 Db 93 TTCTCGCCAGCTTCGCGGCTTTCCCGTCAAGCTCTAAAT 133
 |||||

RESULT 7
 ID ABZ74603
 XX ABZ74603 standard; DNA; 288 BP.
 AC ABZ74603;
 XX
 XX
 DT 12-MAY-2003 (first entry)
 XX
 DE Secreted protein gene 366 genomic fragment HUSGU40, SEQ ID NO:1750.
 XX
 KW Human; secreted protein; cancer; tumour; hyperproliferative disorder;
 KW autoimmune disorder; inflammation; angiogenic diseases; AIDS;
 KW acquired immunodeficiency syndrome; hepatitis; anaemia; wound healing;
 KW drug screening; chromosome identification; chromosome mapping;
 KW cytostatic; gene therapy; antiinflammatory; immunomodulator; anti-HIV;
 KW antianaemic; vulnery; gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200277013-A2.
 XX
 PD 03-OCT-2002.
 XX
 XX 26-MAR-2002; 2002WO-US009370.
 XX
 XX 27-MAR-2001; 2001US-0278650P.
 PR 12-SEP-2001; 2001US-00950082.
 PR 12-SEP-2001; 2001US-00950083.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Rosen CA, Ruben SM;
 PI
 XX WPI; 2003-040578/03.
 DR
 XX New human secreted proteins and nucleic acids, useful for detecting or
 PT treating cancer or other hyperproliferative disorders, autoimmune
 PT disorders, inflammatory disorders, HIV disease, hepatitis or anemia.
 XX
 PS Disclosure; Page 2326; 2474pp; English.
 XX

ABZ73281-ABZ73697 represent cDNAs corresponding to 391 human secreted
 protein genes, and ABP0947-ABP01363 represent the proteins they encode.
 CC ABZ73698-ABZ74687 represent human secreted protein genomic fragments. The
 CC invention also encompasses antibodies specific for the secreted proteins,
 CC the use of the secreted proteins in drug screening and recombinant
 CC vectors and host cells comprising a nucleic acid of the invention. The
 CC secreted proteins are thought to be involved in biological activities
 CC associated with cellular signalling, cellular differentiation, cell
 CC migration, prohormone activation and neurotransmitter activity. The
 CC secreted proteins, nucleic acids encoding them, antibodies or antibody
 CC fragments specific for the secreted proteins, and modulators of protein
 CC activity are useful for diagnosing or treating cancers or other
 CC hyperproliferative disorders. Additionally, the secreted proteins and
 CC their nucleic acids may also be used in the treatment of autoimmune
 CC disorders, inflammatory disorders, diseases involving angiogenesis, AIDS
 CC (acquired immunodeficiency syndrome), hepatitis, anaemia, and to promote
 CC wound healing. Nucleic acids of the invention may be used for chromosome
 CC identification, chromosome mapping, in gene therapy, for identifying
 CC individuals from minute biological samples, as hybridisation probes, and
 CC as molecular weight markers. The present sequence represents a human
 CC secreted protein genomic fragment referred to in the disclosure of the
 CC invention

Query Match 100.0%; Score 101; DB 8; Length 288;
 Best Local Similarity 100.0%; Pred. No. 4.4e-21;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTCCTTCGCTTTCCCTTCCT 60
 |||||
 Db 33 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTCCTTCGCTTTCCCTTCCT 92
 |||||

QY 61 TTCTCGCCAGCTTCGCGGCTTTCCCGTCAAGCTCTAAAT 101
 |||||
 Db 93 TTCTCGCCAGCTTCGCGGCTTTCCCGTCAAGCTCTAAAT 133
 |||||

RESULT 8
 ID ABZ74602
 XX ABZ74602 standard; DNA; 288 BP.
 AC ABZ74602;
 XX
 DT 12-MAY-2003 (first entry)
 XX
 DE Secreted protein gene 366 genomic fragment HUSGU40, SEQ ID NO:1749.
 XX
 KW Human; secreted protein; cancer; tumour; hyperproliferative disorder;
 KW autoimmune disorder; inflammation; angiogenic diseases; AIDS;
 KW acquired immunodeficiency syndrome; hepatitis; anaemia; wound healing;
 KW drug screening; chromosome identification; chromosome mapping;
 KW cytostatic; gene therapy; antiinflammatory; immunomodulator; anti-HIV;
 KW antianaemic; vulnery; gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200277013-A2.
 XX
 PD 03-OCT-2002.
 XX
 XX 26-MAR-2002; 2002WO-US009370.
 XX
 XX 27-MAR-2001; 2001US-0278650P.
 PR 12-SEP-2001; 2001US-00950082.
 PR 12-SEP-2001; 2001US-00950083.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Rosen CA, Ruben SM;
 PI
 XX WPI; 2003-040578/03.
 DR
 XX New human secreted proteins and nucleic acids, useful for detecting or
 PT treating cancer or other hyperproliferative disorders, autoimmune
 PT disorders, inflammatory disorders, HIV disease, hepatitis or anemia.
 XX
 PS Disclosure; Page 2326; 2474pp; English.
 XX

ABZ73281-ABZ73697 represent cDNAs corresponding to 391 human secreted
 protein genes, and ABP0947-ABP01363 represent the proteins they encode.
 CC ABZ73698-ABZ74687 represent human secreted protein genomic fragments. The
 CC invention also encompasses antibodies specific for the secreted proteins,
 CC the use of the secreted proteins in drug screening and recombinant
 CC vectors and host cells comprising a nucleic acid of the invention. The
 CC secreted proteins are thought to be involved in biological activities
 CC associated with cellular signalling, cellular differentiation, cell
 CC migration, prohormone activation and neurotransmitter activity. The
 CC secreted proteins, nucleic acids encoding them, antibodies or antibody
 CC fragments specific for the secreted proteins, and modulators of protein
 CC activity are useful for diagnosing or treating cancers or other
 CC hyperproliferative disorders. Additionally, the secreted proteins and
 CC their nucleic acids may also be used in the treatment of autoimmune
 CC disorders, inflammatory disorders, diseases involving angiogenesis, AIDS
 CC (acquired immunodeficiency syndrome), hepatitis, anaemia, and to promote
 CC wound healing. Nucleic acids of the invention may be used for chromosome
 CC identification, chromosome mapping, in gene therapy, for identifying
 CC individuals from minute biological samples, as hybridisation probes, and
 CC as molecular weight markers. The present sequence represents a human
 CC secreted protein genomic fragment referred to in the disclosure of the
 CC invention


```
CC invention
SQ Sequence 288 BP; 43 A; 87 C; 73 G; 85 T; 0 U; 0 Other;
  Query Match      100.0%; Score 101; DB 8; Length 288;
  Best Local Similarity 100.0%; Pred. No. 4.4e-21;
  Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

  Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCGCTTCCCGTCAAGCTCTAAAT 101
  Db 33 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCGCTTCCCGTCTTCCCTTCCCT 60
  Qy 61 TTCTCGCCACGTTCCGCGGCTTCCCGTCAAGCTCTAAAT 101
  Db 93 TTCTCGCCACGTTCCGCGGCTTCCCGTCAAGCTCTAAAT 133

RESULT 9
ADA98980
ID ADA98980 standard; DNA; 288 BP.
AC ADA98980;
XX
XX
DT 20-NOV-2003 (first entry)
DE
DE Human secreted protein-related DNA sequence #573.
KW human; secreted protein; cardiovascular disorder; arrhythmia;
KW atherosclerosis; stroke; endocarditis; congestive heart failure;
KW rheumatic heart disease; cardiomyopathy; hemorrhoids; varicose veins;
KW migraine; thrombosis; neural disorder; immune system disorder;
KW muscular disorder; reproductive disorder; gastrointestinal disorder;
KW pulmonary disorder; renal disorder; proliferative disorder; cancer; ds.
OS Homo sapiens.
XX
XX WO2003004623-A2.
XX
XX 16-JAN-2003.
XX
XX 26-MAR-2002; 2002WO-US009922.
XX
XX 27-MAR-2001; 2001US-0278650P.
XX
XX 12-SEP-2001; 2001US-00950082.
XX
XX 12-SEP-2001; 2001US-00950083.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
XX
XX WPI; 2003-247946/24.
XX
XX New human secreted polypeptide and nucleic acid molecules, useful for
XX diagnosing, preventing, prognosticating or treating cardiovascular
XX disorders (e.g. arrhythmia, atherosclerosis, cardiomyopathy, or
XX thrombosis).
XX
XX Disclosure; SEQ ID NO 1089; 1572pp; English.
XX
XX The invention comprises the amino acid and coding sequence of human
XX secreted proteins. The DNA and protein sequences of the invention are
XX useful in the treatment of cardiovascular disorders, such as: arrhythmia,
XX atherosclerosis, stroke, endocarditis, congestive heart failure,
XX rheumatic heart disease, cardiomyopathy, hemorrhoids, varicose veins,
XX migraine, or thrombosis. The DNA and protein sequences may also be used
XX for treating or preventing: neural disorders, immune system disorders,
XX muscular disorders, reproductive disorders, gastrointestinal disorders,
XX pulmonary disorders, renal disorders, proliferative disorders and/or
XX cancerous diseases. The present DNA sequence is used in the
XX exemplification of the invention. NOTE: The present sequence is shown on
XX the WIPO website.
XX
XX Sequence 288 BP; 43 A; 87 C; 73 G; 85 T; 0 U; 0 Other;

Query Match      100.0%; Score 101; DB 8; Length 288;
Best Local Similarity 100.0%; Pred. No. 4.4e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCGCTTCCCGTCTTCCCTTCCCT 60
Db 33 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCGCTTCCCGTCTTCCCTTCCCT 92
Qy 61 TTCTCGCCACGTTCCGCGGCTTCCCGTCAAGCTCTAAAT 101
Db 93 TTCTCGCCACGTTCCGCGGCTTCCCGTCAAGCTCTAAAT 133

RESULT 10
ADA98979
ID ADA98979 standard; DNA; 288 BP.
XX
XX ADA98979;
XX
XX 20-NOV-2003 (first entry)
XX
XX Human secreted protein-related DNA sequence #572.
XX
XX human; secreted protein; cardiovascular disorder; arrhythmia;
XX atherosclerosis; stroke; endocarditis; congestive heart failure;
XX rheumatic heart disease; cardiomyopathy; hemorrhoids; varicose veins;
XX migraine; thrombosis; neural disorder; immune system disorder;
XX muscular disorder; reproductive disorder; gastrointestinal disorder;
XX pulmonary disorder; renal disorder; proliferative disorder; cancer; ds.
XX
XX Homo sapiens.
XX
XX WO2003004623-A2.
XX
XX 16-JAN-2003.
XX
XX 26-MAR-2002; 2002WO-US009922.
XX
XX 27-MAR-2001; 2001US-0278650P.
XX
XX 12-SEP-2001; 2001US-00950082.
XX
XX 12-SEP-2001; 2001US-00950083.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
XX
XX WPI; 2003-247946/24.
XX
XX New human secreted polypeptide and nucleic acid molecules, useful for
XX diagnosing, preventing, prognosticating or treating cardiovascular
XX disorders (e.g. arrhythmia, atherosclerosis, cardiomyopathy, or
XX thrombosis).
XX
XX Disclosure; SEQ ID NO 1088; 1572pp; English.
XX
XX The invention comprises the amino acid and coding sequence of human
XX secreted proteins. The DNA and protein sequences of the invention are
XX useful in the treatment of cardiovascular disorders, such as: arrhythmia,
XX atherosclerosis, stroke, endocarditis, congestive heart failure,
XX rheumatic heart disease, cardiomyopathy, hemorrhoids, varicose veins,
XX migraine, or thrombosis. The DNA and protein sequences may also be used
XX for treating or preventing: neural disorders, immune system disorders,
XX muscular disorders, reproductive disorders, gastrointestinal disorders,
XX pulmonary disorders, renal disorders, proliferative disorders and/or
XX cancerous diseases. The present DNA sequence is used in the
XX exemplification of the invention. NOTE: The present sequence is shown on
XX the WIPO website.
XX
XX Sequence 288 BP; 43 A; 87 C; 73 G; 85 T; 0 U; 0 Other;

Query Match      100.0%; Score 101; DB 8; Length 288;
Best Local Similarity 100.0%; Pred. No. 4.4e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCGCTTCCCGTCTTCCCTTCCCT 60
Db 33 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCGCTTCCCGTCTTCCCTTCCCT 92
Qy 61 TTCTCGCCACGTTCCGCGGCTTCCCGTCAAGCTCTAAAT 101
Db 93 TTCTCGCCACGTTCCGCGGCTTCCCGTCAAGCTCTAAAT 133
```

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTCCTTTGCGCTTTCTTCCCTTCT 60
Db |||||
33 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTCCTTTGCGCTTTCTTCCCTTCT 92
Qy 61 TTCTCGCACGTTTCGCGGCTTTCCCGTCAAGCTCTAAAT 101
Db |||||
93 TTCTCGCACGTTTCGCGGCTTTCCCGTCAAGCTCTAAAT 133

RESULT 11

ADA44489
ID ADA44489 standard; DNA; 288 BP.

AC ADA44489;

XX 20-NOV-2003 (first entry)
XX
XX Human secreted protein DNA SEQ ID 682.

DE Gene therapy; human; Antidiabetic; Anorectic; Ophthalmological;

XX Neuroprotective; Cerebroprotective; Antianemic; ds.
KW Homo sapiens.
OS
XX WO2003000865-A2.
PN
XX 03-JAN-2003.
XX
XX 26-MAR-2002; 2002WO-US009105.

XX 27-MAR-2001; 2001US-0278650P.

PR 12-SEP-2001; 2001US-00950082.

PR 12-SEP-2001; 2001US-00950083.

XX (HUMA-) HUMAN GENOME SCI INC.

PA Rosen CA, Ruben SM;

XX WPI; 2003-184045/18.

XX A human secreted protein and nucleic acids useful for preparing a

PT diagnostic or pharmaceutical composition for diagnosing or treating
PT diabetes or conditions related to diabetes, e.g. hyperglycemia, obesity,
PT retinopathy, neuropathy.

XX Disclosure; SEQ ID NO 682; 701pp; English.

XX The invention relates to novel genes and their fragments which are useful
CC for preventing, treating or ameliorating medical conditions e.g. by
CC protein or gene therapy. The genes are isolated from a range of human
CC tissues disclosed in the specification. The nucleic acids and proteins
CC are useful in the diagnosis, treatment and prevention of conditions
CC related to diabetes, e.g. hyperglycaemia, obesity, retinopathy,
CC polynuropathy, atherosclerosis, anaemia, stroke, gangrene, impotence,
CC infection, cataract, renal disorders, or endocrine disorders. The present
CC sequence was used to illustrate the invention.

XX Sequence 288 BP; 43 A; 87 C; 73 G; 85 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 8; Length 288;
Best Local Similarity 100.0%; Pred. No. 4.4e-21;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTCCTTTGCGCTTTCTTCCCTTCT 60
Db |||||
33 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTCCTTTGCGCTTTCTTCCCTTCT 92

Qy 61 TTCTCGCACGTTTCGCGGCTTTCCCGTCAAGCTCTAAAT 101

Db |||||
93 TTCTCGCACGTTTCGCGGCTTTCCCGTCAAGCTCTAAAT 133

RESULT 12

ADA44488
ID ADA44488 standard; DNA; 288 BP.

XX ADA44488;

XX 20-NOV-2003 (first entry)
XX
XX Human secreted protein DNA SEQ ID 681.

XX Gene therapy; human; Antidiabetic; Anorectic; Ophthalmological;
KW Neuroprotective; Cerebroprotective; Antianemic; ds.
XX
XX Homo sapiens.

XX WO2003000865-A2.

PN 03-JAN-2003.

XX 26-MAR-2002; 2002WO-US009105.

XX 27-MAR-2001; 2001US-0278650P.

PR 12-SEP-2001; 2001US-00950082.

PR 12-SEP-2001; 2001US-00950083.

XX (HUMA-) HUMAN GENOME SCI INC.

PA Rosen CA, Ruben SM;

XX WPI; 2003-184045/18.

XX A human secreted protein and nucleic acids useful for preparing a
PT diagnostic or pharmaceutical composition for diagnosing or treating
PT diabetes or conditions related to diabetes, e.g. hyperglycemia, obesity,
PT retinopathy, neuropathy.

XX Disclosure; SEQ ID NO 681; 701pp; English.

XX The invention relates to novel genes and their fragments which are useful
CC for preventing, treating or ameliorating medical conditions e.g. by
CC protein or gene therapy. The genes are isolated from a range of human
CC tissues disclosed in the specification. The nucleic acids and proteins
CC are useful in the diagnosis, treatment and prevention of conditions
CC related to diabetes, e.g. hyperglycaemia, obesity, retinopathy,
CC polynuropathy, atherosclerosis, anaemia, stroke, gangrene, impotence,
CC infection, cataract, renal disorders, or endocrine disorders. The present
CC sequence was used to illustrate the invention.

XX Sequence 288 BP; 43 A; 87 C; 73 G; 85 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 8; Length 288;
Best Local Similarity 100.0%; Pred. No. 4.4e-21;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTCCTTTGCGCTTTCTTCCCTTCT 60
Db |||||
33 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTCCTTTGCGCTTTCTTCCCTTCT 92

Qy 61 TTCTCGCACGTTTCGCGGCTTTCCCGTCAAGCTCTAAAT 101

Db |||||
93 TTCTCGCACGTTTCGCGGCTTTCCCGTCAAGCTCTAAAT 133

RESULT 13

ABZ68124
ID ABZ68124 standard; DNA; 288 BP.

XX ABZ68124;

XX 26-MAR-2003 (first entry)
XX
XX Human secreted protein encoding genomic DNA SEQ ID NO 1647.

XX Human; secreted protein; nootropic; neuroprotective; cytostatic;
KW virucide; dermatological; immunosuppressive; antiinflammatory; anti-HIV;
KW vulnary; antibacterial; antiparkinsonian; antiskickling; antianaemic;
KW antiarthritic; cancer; antirheumatic; hepatotropic; cerebroprotective;
KW antiinflammatory; antiallergic; antidiabetic; antitumor; anticonvulsant;
KW antifungal; antiparasitic; cardiant; immune disorder; infection; vaccine;
KW cardiovascular disorder; neurological disease; nephrotropic;
KW gene therapy; gene; ds.
XX Homo sapiens.
XX WO200277186-A2.
XX 03-OCT-2002.
XX 26-MAR-2002; 2002WO-US009188.
XX 27-MAR-2001; 2001US-0278650P.
PR 12-SEP-2001; 2001US-00950082.
PR 12-SEP-2001; 2001US-00950083.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Ruben SM;
PI WPI; 2003-040583/03.
XX New human secreted proteins encoded by genes contained in cDNA clones
PT (e.g. HGAC19), useful for preventing, treating or diagnosing e.g. AIDS,
PT multiple sclerosis, herpes virus, leukemia, tick-borne encephalitis or
PT West Nile fever.
XX Disclosure; Page 2268; 2423pp; English.
XX The invention relates to novel human genes (ABZ66891-ABZ68209) and the
CC encoded secreted proteins (ABP99470-ABP99872) useful for preventing,
CC treating or ameliorating medical conditions e.g. by protein or gene
CC therapy. The genes are isolated from a range of human tissues disclosed
CC in the specification. The nucleic acids, proteins, antibodies and
CC (ant)agonists are useful in the diagnosis, treatment and prevention of:
CC (a) cancer, e.g. breast and ovarian cancer and other cancers of the
CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
CC lung or urogenital; (b) immune disorders e.g. Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; (c) cardiovascular disorders such as
CC myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g.
CC cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
CC bacterial, fungal and parasitic infections
SQ Sequence 288 BP; 43 A; 87 C; 73 G; 85 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 10; Length 288;
Best Local Similarity 100.0%; Pred. No. 4.4e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCGTGACCGGTACACTTGCAGCGCCCTAGCGCGGCTTCCTTTCCTTCCTTCCT 60
DB 33 GCGTGACCGGTACACTTGCAGCGCCCTAGCGCGGCTTCCTTTCCTTCCTTCCT 92
QY 61 TTCTCGCCACGTTGCGCGGCTTTCCCGTCAAGCTCTAAAT 101
DB 93 TTCTCGCCACGTTGCGCGGCTTTCCCGTCAAGCTCTAAAT 133
RESULT 14
ID ABZ68123
XX AC ABZ68123;
XX 26-MAR-2003 (first entry)
DT

XX Human secreted protein encoding genomic DNA SEQ ID NO 1646.
DE
XX Human; secreted protein; nootropic; neuroprotective; cytostatic;
KW virucide; dermatological; immunosuppressive; antiinflammatory; anti-HIV;
KW vulnary; antibacterial; antiparkinsonian; antiskickling; antianaemic;
KW antiarthritic; cancer; antirheumatic; hepatotropic; cerebroprotective;
KW antiinflammatory; antiallergic; antidiabetic; antitumor; anticonvulsant;
KW antifungal; antiparasitic; cardiant; immune disorder; infection; vaccine;
KW cardiovascular disorder; neurological disease; nephrotropic;
KW gene therapy; gene; ds.
XX Homo sapiens.
XX WO200277186-A2.
XX 03-OCT-2002.
XX 26-MAR-2002; 2002WO-US009188.
XX 27-MAR-2001; 2001US-0278650P.
PR 12-SEP-2001; 2001US-00950082.
PR 12-SEP-2001; 2001US-00950083.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Ruben SM;
PI WPI; 2003-040583/03.
XX New human secreted proteins encoded by genes contained in cDNA clones
PT (e.g. HGAC19), useful for preventing, treating or diagnosing e.g. AIDS,
PT multiple sclerosis, herpes virus, leukemia, tick-borne encephalitis or
PT West Nile fever.
XX Disclosure; Page 2268; 2423pp; English.
XX The invention relates to novel human genes (ABZ66891-ABZ68209) and the
CC encoded secreted proteins (ABP99470-ABP99872) useful for preventing,
CC treating or ameliorating medical conditions e.g. by protein or gene
CC therapy. The genes are isolated from a range of human tissues disclosed
CC in the specification. The nucleic acids, proteins, antibodies and
CC (ant)agonists are useful in the diagnosis, treatment and prevention of:
CC (a) cancer, e.g. breast and ovarian cancer and other cancers of the
CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
CC lung or urogenital; (b) immune disorders e.g. Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; (c) cardiovascular disorders such as
CC myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g.
CC cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
CC bacterial, fungal and parasitic infections
SQ Sequence 288 BP; 43 A; 87 C; 73 G; 85 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 10; Length 288;
Best Local Similarity 100.0%; Pred. No. 4.4e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCGTGACCGGTACACTTGCAGCGCCCTAGCGCGGCTTCCTTTCCTTCCTTCCT 60
DB 33 GCGTGACCGGTACACTTGCAGCGCCCTAGCGCGGCTTCCTTTCCTTCCTTCCT 92
QY 61 TTCTCGCCACGTTGCGCGGCTTTCCCGTCAAGCTCTAAAT 101
DB 93 TTCTCGCCACGTTGCGCGGCTTTCCCGTCAAGCTCTAAAT 133
RESULT 15
ID ABZ68010
XX AC ABZ68010;
XX 26-MAR-2003 (first entry)
DT

XX 26-MAR-2003 (first entry)
DT Human secreted protein encoding genomic DNA SEQ ID NO 1533.
DE
DE
XX Human; secreted protein; nootropic; neuroprotective; cytostatic;
XX virucide; dermatological; immunosuppressive; antiinflammatory; anti-HIV;
KW vulnery; antibacterial; antiparkinsonian; antistickling; antianaemic;
KW antiarthritic; cancer; antirheumatic; hepatotropic; cerebroprotective;
KW antinflammatory; antiallergic; antidiabetic; antilucer; anticonvulsant;
KW antifungal; antiparasitic; cardiant; immune disorder; infection; vaccine;
KW cardiovascular disorder; neurological disease; nephrotropic;
KW gene therapy; gene; ds.
XX
OS Homo sapiens.
XX
XX WO200277186-A2.
XX
XX 03-OCT-2002.
XX
XX 26-MAR-2002; 2002WO-US009188.
XX
XX 27-MAR-2001; 2001US-0278650P.
PR 12-SEP-2001; 2001US-00950082.
PR 12-SEP-2001; 2001US-00950083.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
FA
XX Rosen CA, Ruben SM;
XX WPI; 2003-040583/03.
XX
XX New human secreted proteins encoded by genes contained in cDNA clones
PT (e.g. HGAC19), useful for preventing, treating or diagnosing e.g. AIDS,
PT multiple sclerosis, herpes virus, leukemia, tick-borne encephalitis or
PT West Nile fever.
XX
XX Disclosure; Page 2154; 2423pp; English.
PS
XX The invention relates to novel human genes (ABZ66891-ABZ68209) and the
CC encoded secreted proteins (ABP99470-ABP99872) useful for preventing,
CC treating or ameliorating medical conditions e.g. by protein or gene
CC therapy. The genes are isolated from a range of human tissues disclosed
CC in the specification. The nucleic acids, proteins, antibodies and
CC (ant)agonists are useful in the diagnosis, treatment and prevention of:
CC (a) cancer, e.g. breast and ovarian cancer and other cancers of the
CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
CC lung or urogenital; (b) immune disorders e.g. Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; (c) cardiovascular disorders such as
CC myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g.
CC cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
CC bacterial, fungal and parasitic infections
XX
SQ Sequence 288 BP; 43 A; 87 C; 73 G; 85 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 10; Length 288;
Best Local Similarity 100.0%; Pred. No. 4.4e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCGTGACCGCTACACTGCGACGCGCCCTAGCGCCGCTCTTCGCTTCTTCCTCCT 60
Dy 33 GCGTGACCGCTACACTGCGACGCGCCCTAGCGCCGCTCTTCGCTTCTTCCTCCT 92
Qy 61 TTCTCGCCACGTTCCGCGGCTTTCCCGTCAAGCTCTAAAT 101
Dy 93 TTCTCGCCACGTTCCGCGGCTTTCCCGTCAAGCTCTAAAT 133

Search completed: July 14, 2005, 07:01:48
Job time : 145.448 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 05:15:57 ; Search time 961.667 Seconds
(without alignments)
3997.736 Million cell updates/sec

Title: US-09-482-682-64_COPY_3565_3665

Perfect score: 101

Sequence: 1 gcgtgaccgtacacttgcc.....ttccccgcagctcttaaat 101

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST.*

1: gb_est1.*
2: gb_est2.*
3: gb_hic.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gss1.*
9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101	100.0	113	8	AQ265696
2	101	100.0	116	8	AQ263357
3	101	100.0	142	8	AQ265335
4	101	100.0	150	8	AQ264698
5	101	100.0	158	8	AQ026683
6	101	100.0	162	8	AQ061259
7	101	100.0	163	8	AQ196514
8	101	100.0	164	4	BW812626
9	101	100.0	166	8	AQ112846
10	101	100.0	168	4	BW812644
11	101	100.0	173	8	AQ310689
12	101	100.0	192	8	AQ280364
13	101	100.0	200	8	AQ278442
14	101	100.0	211	8	AQ280347
15	101	100.0	217	4	B1677411
16	101	100.0	218	8	AQ280220
17	101	100.0	224	4	BW517095
18	101	100.0	230	8	AQ279924
19	101	100.0	236	6	CD282304
20	101	100.0	249	4	B1677389
21	101	100.0	255	4	B1940501
22	101	100.0	261	8	AQ310638
23	101	100.0	264	6	CD280795
24	101	100.0	271	8	B83638

25	101	100.0	271	8	AQ427025
26	101	100.0	275	8	AZ212954
27	101	100.0	276	6	CB867747
28	101	100.0	277	8	AQ263549
29	101	100.0	281	2	BF542624
30	101	100.0	286	7	CK884250
31	101	100.0	286	8	AQ074694
32	101	100.0	288	5	BW573920
33	101	100.0	292	8	B78686
34	101	100.0	307	8	B78705
35	101	100.0	309	7	CK886580
36	101	100.0	309	7	CV162119
37	101	100.0	310	8	AQ009663
38	101	100.0	312	2	BF524055
39	101	100.0	324	8	AQ008830
40	101	100.0	324	8	AQ017932
41	101	100.0	325	6	CD279858
42	101	100.0	325	6	CD281090
43	101	100.0	327	6	CD282801
44	101	100.0	327	8	AQ026530
45	101	100.0	332	8	AQ009332

ALIGNMENTS

RESULT 1
AQ265696
LOCUS
DEFINITION
CITBI-El-2503El.TR CITBI-El Homo sapiens genomic clone 2503El,
genomic survey sequence.
ACCESSION
AQ265696
VERSION
AQ265696.1 GI:3791450
KEYWORDS
GSS.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 (bases 1 to 113)
Adams M.D., Rounsley, S.D., Zhao, S., Bass, S., Linher, K., Golden, K.,
Barry, K., Granger, D., Suh, E., Wible, C., Shizuya, H., Simon, M. and
Venter, J.C.
Use of a random human BAC End Sequence Database for Sequence-Ready
Map Building
Unpublished (1998)
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.

TITLE
JOURNAL
COMMENT

FEATURES
source
1. .113
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="2503El"
/sex="male"
/cell_type="sperm"
/clone_lib="CITBI-El"
/note="Vector: pBelO8A11; Site_1: EcoRI; Site_2: EcoRI;
CalTech Human BAC Library D"

Query Match 100.0%; Score 101; DB 8; Length 113;
Best Local Similarity 100.0%; Pred. No. 9.1e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCCCTTTTCGCTTCTTCCCTTCCT 60
Db 8 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCCCTTTTCGCTTCTTCCCTTCCT 67

Qy 61 TTCTCGCCACGTTGCGCGGCTTCCCGCTCAAGCTCTAAAT 101
Db 68 TTCTCGCCACGTTGCGCGGCTTCCCGCTCAAGCTCTAAAT 108

RESULT 2
AQ263357 116 bp DNA linear GSS 27-OCT-1998
LOCUS CITBI-EI-2503C21.TR CITBI-EI Homo sapiens genomic clone 2503C21,
DEFINITION genomic survey sequence.
ACCESSION AQ263357
VERSION AQ263357.1 GI:3790953
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 116)
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Baas,S., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
Venter,J.C.
Use of a random human BAC End Sequence Database for Sequence-Ready
Map Building
JOURNAL Unpublished (1998)
COMMENT Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.
FEATURES             Location/Qualifiers
     source           1..116
     organism="Homo sapiens"
     mol_type="genomic DNA"
     db_xref="taxon:9606"
     clone="2503C21"
     sex="male"
     cell_type="sperm"
     clone_lib="CITBI-EI"
     note="Vector: pBelobAC11; Site_1: EcoRI; Site_2: EcoRI;
CalTech Human BAC Library D"
ORIGIN
Query Match 100.0%; Score 101; DB 8; Length 116;
Best Local Similarity 100.0%; Pred. No. 9.2e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCCCTTTTCGCTTCTTCCCTTCCT 60
Db 8 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCCCTTTTCGCTTCTTCCCTTCCT 67

Qy 61 TTCTCGCCACGTTGCGCGGCTTCCCGCTCAAGCTCTAAAT 101
Db 68 TTCTCGCCACGTTGCGCGGCTTCCCGCTCAAGCTCTAAAT 108

RESULT 3
AQ265335 142 bp DNA linear GSS 27-OCT-1998
LOCUS CITBI-EI-2508M1.TR CITBI-EI Homo sapiens genomic clone 2508M1,
DEFINITION genomic survey sequence.
ACCESSION AQ265335
VERSION AQ265335.1 GI:3793535
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 142)
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Baas,S., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
Venter,J.C.
Use of a random human BAC End Sequence Database for Sequence-Ready
Map Building
JOURNAL Unpublished (1998)
COMMENT Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.
FEATURES             Location/Qualifiers
     source           1..142
     organism="Homo sapiens"
     mol_type="genomic DNA"
     db_xref="taxon:9606"
     clone="2508M1"
     sex="male"
     cell_type="sperm"
     clone_lib="CITBI-EI"
     note="Vector: pBelobAC11; Site_1: EcoRI; Site_2: EcoRI;
CalTech Human BAC Library D"
ORIGIN
Query Match 100.0%; Score 101; DB 8; Length 142;
Best Local Similarity 100.0%; Pred. No. 9.3e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCCCTTTTCGCTTCTTCCCTTCCT 60
Db 8 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCCCTTTTCGCTTCTTCCCTTCCT 67

Qy 61 TTCTCGCCACGTTGCGCGGCTTCCCGCTCAAGCTCTAAAT 101
Db 68 TTCTCGCCACGTTGCGCGGCTTCCCGCTCAAGCTCTAAAT 108

RESULT 4
AQ264698 150 bp DNA linear GSS 27-OCT-1998
LOCUS CITBI-EI-2508E1.TR CITBI-EI Homo sapiens genomic clone 2508E1,
DEFINITION genomic survey sequence.
ACCESSION AQ264698
VERSION AQ264698.1 GI:3792898
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 150)
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Baas,S., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
Venter,J.C.
Use of a random human BAC End Sequence Database for Sequence-Ready
Map Building
JOURNAL Unpublished (1998)
COMMENT Contact: Mark Adams
Department of Eukaryotic Genomics

```

```

ACCESSION AQ265335
VERSION AQ265335.1 GI:3793535
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 142)
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Baas,S., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
Venter,J.C.
Use of a random human BAC End Sequence Database for Sequence-Ready
Map Building
JOURNAL Unpublished (1998)
COMMENT Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.
FEATURES             Location/Qualifiers
     source           1..142
     organism="Homo sapiens"
     mol_type="genomic DNA"
     db_xref="taxon:9606"
     clone="2508M1"
     sex="male"
     cell_type="sperm"
     clone_lib="CITBI-EI"
     note="Vector: pBelobAC11; Site_1: EcoRI; Site_2: EcoRI;
CalTech Human BAC Library D"
ORIGIN
Query Match 100.0%; Score 101; DB 8; Length 142;
Best Local Similarity 100.0%; Pred. No. 9.3e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCCCTTTTCGCTTCTTCCCTTCCT 60
Db 8 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCCCTTTTCGCTTCTTCCCTTCCT 67

Qy 61 TTCTCGCCACGTTGCGCGGCTTCCCGCTCAAGCTCTAAAT 101
Db 68 TTCTCGCCACGTTGCGCGGCTTCCCGCTCAAGCTCTAAAT 108

RESULT 4
AQ264698 150 bp DNA linear GSS 27-OCT-1998
LOCUS CITBI-EI-2508E1.TR CITBI-EI Homo sapiens genomic clone 2508E1,
DEFINITION genomic survey sequence.
ACCESSION AQ264698
VERSION AQ264698.1 GI:3792898
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 150)
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Baas,S., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
Venter,J.C.
Use of a random human BAC End Sequence Database for Sequence-Ready
Map Building
JOURNAL Unpublished (1998)
COMMENT Contact: Mark Adams
Department of Eukaryotic Genomics

```

The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org

Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tldb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.

FEATURES

Location/Qualifiers
1..150

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="2508E1"
/sex="male"
/cell_type="sperm"
/clone_lib="CITBi-E1"
/note="Vector: pBelOBAC11; Site 1: EcoRI; Site 2: EcoRI;
CalTech Human BAC Library D"

ORIGIN

Query Match 100.0%; Score 101; DB 8; Length 150;
Best Local Similarity 100.0%; Pred. No. 9.3e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGGTACACTTGCAGCGCCTAGCGCGCTCTTTCGCTTCTTCCCTTCCT 60
Db 8 GCGTGACCGGTACACTTGCAGCGCCTAGCGCGCTCTTTCGCTTCTTCCCTTCCT 67
Qy 61 TTCTCGCCACGTTTCGGCGCTTCCCGTCAAGCTCTAAAT 101
Db 68 TTCTCGCCACGTTTCGGCGCTTCCCGTCAAGCTCTAAAT 108

RESULT 5

LOCUS AQ026683 158 bp DNA linear GSS 30-JUN-1998
DEFINITION CIT-HSP-2314E2.TR CIT-HSP Homo sapiens genomic clone 2314E2,
genomic survey sequence.

ACCESSION AQ026683
VERSION AQ026683.1 GI:3266905
KEYWORDS GSS.
SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 158)
Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.

Use of a random BAC End Sequence Database for Sequence-Ready Map

Building (1998)

JOURNAL

Unpublished (1998)

Contact: Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: mdadams@tigr.org

Clones are available from Research Genetics (info@resgen.com). BAC

end search page:

http://www.tigr.org/tldb/humgen/bac_end_search/bac_end_search.html.

Seq primer: M13 Reverse

Class: BAC ends.

FEATURES

Location/Qualifiers
1..158

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="2314E2"

/sex="Male"
/cell_type="Sperm"
/clone_lib="CIT-HSP"
/note="Vector: pBelOBAC11; Site 1: HindIII; Site 2:
HindIII"

ORIGIN

Query Match 100.0%; Score 101; DB 8; Length 158;
Best Local Similarity 100.0%; Pred. No. 9.4e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGGTACACTTGCAGCGCCTAGCGCGCTCTTTCGCTTCTTCCCTTCCT 60
Db 52 GCGTGACCGGTACACTTGCAGCGCCTAGCGCGCTCTTTCGCTTCTTCCCTTCCT 111
Qy 61 TTCTCGCCACGTTTCGGCGCTTCCCGTCAAGCTCTAAAT 101
Db 112 TTCTCGCCACGTTTCGGCGCTTCCCGTCAAGCTCTAAAT 152

RESULT 6

LOCUS AQ061259 162 bp DNA linear GSS 31-JUL-1998
DEFINITION CIT-HSP-2352N1.TR CIT-HSP Homo sapiens genomic clone 2352N1,
genomic survey sequence.

ACCESSION AQ061259

VERSION AQ061259.1 GI:3363171

KEYWORDS GSS.

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 162)

Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.

Use of a random BAC End Sequence Database for Sequence-Ready Map

Building (1998)

Unpublished (1998)

Contact: Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: mdadams@tigr.org

Clones are available from Research Genetics (info@resgen.com). BAC

end search page:

http://www.tigr.org/tldb/humgen/bac_end_search/bac_end_search.html.

Seq primer: M13 Reverse

Class: BAC ends.

Location/Qualifiers

1..162

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/clone="2352N1"

/sex="Male"

/cell_type="Sperm"

/clone_lib="CIT-HSP"

/note="Vector: pBelOBAC11; Site 1: HindIII; Site 2:
HindIII"

ORIGIN

Query Match 100.0%; Score 101; DB 8; Length 162;
Best Local Similarity 100.0%; Pred. No. 9.4e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGGTACACTTGCAGCGCCTAGCGCGCTCTTTCGCTTCTTCCCTTCCT 60
Db 52 GCGTGACCGGTACACTTGCAGCGCCTAGCGCGCTCTTTCGCTTCTTCCCTTCCT 111
Qy 61 TTCTCGCCACGTTTCGGCGCTTCCCGTCAAGCTCTAAAT 101

```

Db      112 TTCTCGCACGCTTCCGCGCTTTCCCGCTCAAGCTCTAAAT 152
|||||
RESULT 7
AQ196514
LOCUS   163 bp    DNA    linear    GSS 16-SEP-1998
DEFINITION   CIT-HSP-2385C23.TR CIT-HSP Homo sapiens genomic clone 2385C23,
              genomic survey sequence.
ACCESSION   AQ196514
VERSION     AQ196514.1 GI:3603876
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 163)
AUTHORS    Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
            Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
            Venter,J.C.
TITLE      Use of a random human BAC End Sequence Database for Sequence-Ready
            Map Building
JOURNAL    Unpublished (1998)
COMMENT     Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13 Reverse
            Class: BAC ends.
FEATURES             source
     source           1..163
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="2385C23"
                     /sex="Male"
                     /cell_type="Sperm"
                     /clone_lib="CIT-HSP"
                     /note="Vector: pBelOBAC11; Site_1: HindIII; Site_2:
                     HindIII"
ORIGIN
Query Match      100.0%; Score 101; DB 8; Length 163;
Best Local Similarity 100.0%; Pred. No. 9.4e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTTCTTCGCTTCTTCCTTCCT 60
        |||||||
Db      52 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTTCTTCGCTTCTTCCTTCCT 111
        |||||||

Qy      61 TTCTCGCACGCTTCCGCGCTTTCCCGCTCAAGCTCTAAAT 101
        |||||||
Db      112 TTCTCGCACGCTTCCGCGCTTTCCCGCTCAAGCTCTAAAT 152
        |||||||

RESULT 8
BM812626
LOCUS   164 bp    mRNA    linear    EST 05-MAR-2002
DEFINITION   rt03b06.v2 Pristionchus pacificus mixed stage SL1 TOPO v1 Murphy
            Chiapelli McCarter Pristionchus pacificus cDNA 5', mRNA sequence.
ACCESSION   BM812626
VERSION     BM812626.1 GI:19148640
KEYWORDS    EST.
SOURCE      Pristionchus pacificus
            Pristionchus pacificus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
            Neodiplogasteridae; Pristionchus.

```

```

REFERENCE
AUTHORS    McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J.,
            Wylie,T., Dante,M., Marra,M., Hillier,L., Bennett,J., Franklin,B.,
            Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C.,
            Tsagarishvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C.,
            Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T.,
            Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
            McCann,R., Waterston,R. and Wilson,R.
            The Washington Univ. Nematode EST Project, 1999
            Unpublished (1999)
            Contact: McCarter JP
            The Washington Univ. Nematode EST Project, 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            The library was constructed by Claire Murphy, Brandi Chiapelli, and
            Dr. James McCarter at Washington University, St. Louis. DNA
            Sequencing by: Washington University Genome Sequencing Center
            Seq primer: -40RP from Gibco
            High quality sequence stop: 161.
FEATURES             source
     source           1..164
                     /organism="Pristionchus pacificus"
                     /mol_type="mRNA"
                     /db_xref="taxon:54126"
                     /dev_stage="Mixed stage"
                     /lab_host="DH10B"
                     /clone_lib="Pristionchus pacificus mixed stage SL1 TOPO v1
                     Murphy Chiapelli McCarter"
                     /note="Vector: pCRII-TOPO (Invitrogen); Site_1: EcoRI;
                     Site_2: EcoRI; The library was constructed by Claire
                     Murphy, Brandi Chiapelli, and Dr. James McCarter at
                     Washington University, St. Louis. Oligo(dT)-SL1 PCR based
                     library. Pristionchus pacificus mixed stage cDNA PCR
                     products of size 3400 nucleotides containing SL1 on the
                     5' end and oligo(dT) on the 3' end were non-directionally
                     cloned into pCRII-TOPO(Invitrogen) following the TOPO TA
                     cloning protocol."
ORIGIN
Query Match      100.0%; Score 101; DB 4; Length 164;
Best Local Similarity 100.0%; Pred. No. 9.4e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTTCTTCGCTTCTTCCTTCCT 60
        |||||||
Db      51 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTTCTTCGCTTCTTCCTTCCT 110
        |||||||

Qy      61 TTCTCGCACGCTTCCGCGCTTTCCCGCTCAAGCTCTAAAT 101
        |||||||
Db      111 TTCTCGCACGCTTCCGCGCTTTCCCGCTCAAGCTCTAAAT 151
        |||||||

RESULT 9
AQ112846
LOCUS   166 bp    DNA    linear    GSS 29-AUG-1998
DEFINITION   CIT-HSP-2376E2.TR CIT-HSP Homo sapiens genomic clone 2376E2,
            genomic survey sequence.
ACCESSION   AQ112846
VERSION     AQ112846.1 GI:3488967
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 166)
AUTHORS    Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
            Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
            Venter,J.C.
            Use of a random human BAC End Sequence Database for Sequence-Ready
            Map Building
TITLE

```


Unpublished (1998)
 Other_GSSs: CIT-HSP-2376E2.TF
 Contact: Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: mdadams@tigr.org
 Clones are available from Research Genetics (info@resgen.com). BAC
 end search page:
 http://www.tigr.org/tcdb/humgen/bac_end_search/bac_end_search.html.
 Seq primer: M13 Reverse
 Class: BAC ends.

FEATURES
 source
 Location/Qualifiers
 1..166
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="2376E2"
 /sex="Male"
 /cell_type="Sperm"
 /clone_lib="CIT-HSP"
 /note="Vector: pBelOBAC11; Site_1: HindIII; Site_2:
 HindIII"

ORIGIN
 Query Match 100.0%; Score 101; DB 8; Length 166;
 Best Local Similarity 100.0%; Pred. No. 9.4e-18;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 GCGTGACCGCTACACTTGCACAGCGCCCTAGCGCGCGCTCTTTCGCTTTCTTCCCTTCTCT 60
 52 GCGTGACCGCTACACTTGCACAGCGCCCTAGCGCGCGCTCTTTCGCTTTCTTCCCTTCTCT 111
 61 TTCTCGCCACGTTTCGCGCGCTTTCCCGCTCAAGCTCTAAAT 101
 112 TTCTCGCCACGTTTCGCGCGCTTTCCCGCTCAAGCTCTAAAT 152

RESULT 10
 EM812644
 LOCUS
 DEFINITION
 Pristionchus pacificus
 Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
 Neodiplogasteridae; Pristionchus.
 1 (bases 1 to 168)
 McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J.,
 Wylie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B.,
 Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C.,
 Tsagarishvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C.,
 Harvey,N.K., Sepcoe,M., Allen,M., Person,B., Swaller,T.,
 Underwood,S., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
 McCam,R., Waterston,R. and Wilson.R.
 The Washington Univ. Nematode EST Project, 1999
 Unpublished (1999)
 Contact: McCarter JP
 The Washington Univ. Nematode EST Project, 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 The library was constructed by Claire Murphy, Brandi Chisapelli, and
 Dr. James McCarter at Washington University, St. Louis. DNA
 Sequencing by: Washington University Genome Sequencing Center
 Seq primer: -40RP from Gbco

```

High quality sequence stop: 167.
FEATURES
    source
        1. .168
            Location/Qualifiers
                /organism="Pristionchus pacificus"
                /mol_type="mRNA"
                /db_xref="taxon:541126"
                /dev_stage="Mixed stage"
                /lab_host="DH108"
                /clone_lib="Pristionchus pacificus mixed stage SL1 TOPO v1"
                /vector="pCRII-TOPO (Invitrogen) : Site 1: EcoRI;
                Site 2: EcoRI: The library was constructed by Claire
                Murphy, Brandi Chipelli, and Dr. James McCarter at
                Washington University, St. Louis. Oligo(dT)-SL1 PCR based
                library. Pristionchus pacificus mixed stage cDNA PCR
                products of size >400 nucleotides containing SL1 on the
                5' end and oligo(dT) on the 3' end were non-directionally
                cloned into pCRII-TOPO(Invitrogen) following the Topo TA
                Cloning protocol."
ORIGIN
    Query Match 100.0%; Score 101; DB 4; Length 168;
    Best Local Similarity 100.0%; Pred No. 9.4e-18;
    Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTCTCTTCGCTTTCTCCCTTCCT 60
    |||
Db 52 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTCTCTTCGCTTTCTCCCTTCCT 111
    |||

QY 61 TTCTCGCCAGGTCGCGCGCTTCCCGTCAAGCTCTAAAT 101
    |||
Db 112 TTCTCGCCAGGTCGCGCGCTTCCCGTCAAGCTCTAAAT 152
    |||

RESULT 11
AQ310689
LOCUS
DEFINITION
    AQ310689 173 bp DNA linear GSS 22-DEC-1998
    CITR1-EI-2520L11.TR CITR1-EI Homo sapiens genomic clone 2520L11,
    genomic survey sequence.
ACCESSION
    AQ310689
VERSION
    AQ310689.1 GI:4042502
KEYWORDS
    GSS.
SOURCE
    Homo sapiens (human)
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
    1 (bases 1 to 173)
    Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
    Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
    Venter,J.C.
    Use of a random human BAC End Sequence Database for Sequence-Ready
    Map Building
    Unpublished (1998)
JOURNAL
    Contact: Shaying Zhao, William Nierman, Mark Adams
    Department of Eukaryotic Genomics
    The Institute for Genomic Research
    9712 Medical Center Dr., Rockville, MD 20850
    Tel: 301 838 0200
    Fax: 301 838 0208
    Email: hbe@tigr.org
    Clones are available from Research Genetics (info@resgen.com). BAC
    end search page:
    http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.
    Seq primer: M13 Reverse
    Class: BAC end.
FEATURES
    Location/Qualifiers
        1. .173
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /clone="2520L11"
            /sex="male"
            /cell_type="sperm"

```

```

/clone lib="CITBI-E1"
/notes=Vector: pBelOBAC11; Site_1: EcoRI; Site_2: EcoRI;
CalTech Human BAC Library D"

ORIGIN
Query Match      100.0%; Score 101; DB 8; Length 173;
Best Local Similarity 100.0%; Pred. No. 9.4e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTCCTTTTCGCTTTCTTCCCTTCCT 60
Db 8 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTCCTTTTCGCTTTCTTCCCTTCCT 67

Qy 61 TTCTCGCCACGCTTCGCCGGCTTTCCCGTCAAGCTCTAAAT 101
Db 68 TTCTCGCCACGCTTCGCCGGCTTTCCCGTCAAGCTCTAAAT 108

RESULT 12
LOCUS      AQ280364
DEFINITION CITBI-E1-2518K13.TR CITBI-E1 Homo sapiens genomic clone 2518K13,
            genomic survey sequence.
ACCESSION  AQ280364
VERSION    AQ280364.1 GI:3906183
KEYWORDS   GSS.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 192)
AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
            Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
            Venter,J.C.
TITLE     Use of a random human BAC End Sequence Database for Sequence-Ready
            Map Building
JOURNAL    Unpublished (1998)
COMMENT    Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13 Reverse
            Class: BAC ends.

FEATURES             Location/Qualifiers
     source            1..192
     map_feature        1..192
     other_GSSs        CITBI-E1-2518K13.TF
     Contact: Mark Adams
     Department of Eukaryotic Genomics
     The Institute for Genomic Research
     9712 Medical Center Dr., Rockville, MD 20850, USA
     Tel: 301 838 0200
     Fax: 301 838 0208
     Email: mdadams@tigr.org
     Clones are available from Research Genetics (info@resgen.com). BAC
     end search page:
     http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
     Seq primer: M13 Reverse
     Class: BAC ends.

FEATURES             Location/Qualifiers
     source            1..192
     map_feature        1..192
     other_GSSs        CITBI-E1-2518K13.TF
     Contact: Mark Adams
     Department of Eukaryotic Genomics
     The Institute for Genomic Research
     9712 Medical Center Dr., Rockville, MD 20850, USA
     Tel: 301 838 0200
     Fax: 301 838 0208
     Email: mdadams@tigr.org
     Clones are available from Research Genetics (info@resgen.com). BAC
     end search page:
     http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
     Seq primer: M13 Reverse
     Class: BAC ends.

ORIGIN
Query Match      100.0%; Score 101; DB 8; Length 192;
Best Local Similarity 100.0%; Pred. No. 9.5e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTCCTTTTCGCTTTCTTCCCTTCCT 60
Db 8 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTCCTTTTCGCTTTCTTCCCTTCCT 67

Qy 61 TTCTCGCCACGCTTCGCCGGCTTTCCCGTCAAGCTCTAAAT 101
Db 68 TTCTCGCCACGCTTCGCCGGCTTTCCCGTCAAGCTCTAAAT 108

RESULT 13
LOCUS      AQ278442
DEFINITION CITBI-E1-2519N1.TR CITBI-E1 Homo sapiens genomic clone 2519N1,
            genomic survey sequence.
ACCESSION  AQ278442
VERSION    AQ278442.1 GI:3904410
KEYWORDS   GSS.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 200)
AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
            Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
            Venter,J.C.
TITLE     Use of a random human BAC End Sequence Database for Sequence-Ready
            Map Building
JOURNAL    Unpublished (1998)
COMMENT    Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13 Reverse
            Class: BAC ends.

FEATURES             Location/Qualifiers
     source            1..200
     map_feature        1..200
     other_GSSs        CITBI-E1-2519N1.TF
     Contact: Mark Adams
     Department of Eukaryotic Genomics
     The Institute for Genomic Research
     9712 Medical Center Dr., Rockville, MD 20850, USA
     Tel: 301 838 0200
     Fax: 301 838 0208
     Email: mdadams@tigr.org
     Clones are available from Research Genetics (info@resgen.com). BAC
     end search page:
     http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
     Seq primer: M13 Reverse
     Class: BAC ends.

FEATURES             Location/Qualifiers
     source            1..200
     map_feature        1..200
     other_GSSs        CITBI-E1-2519N1.TF
     Contact: Mark Adams
     Department of Eukaryotic Genomics
     The Institute for Genomic Research
     9712 Medical Center Dr., Rockville, MD 20850, USA
     Tel: 301 838 0200
     Fax: 301 838 0208
     Email: mdadams@tigr.org
     Clones are available from Research Genetics (info@resgen.com). BAC
     end search page:
     http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
     Seq primer: M13 Reverse
     Class: BAC ends.

ORIGIN
Query Match      100.0%; Score 101; DB 8; Length 200;
Best Local Similarity 100.0%; Pred. No. 9.5e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTCCTTTTCGCTTTCTTCCCTTCCT 60
Db 12 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTCCTTTTCGCTTTCTTCCCTTCCT 71

Qy 61 TTCTCGCCACGCTTCGCCGGCTTTCCCGTCAAGCTCTAAAT 101
Db 72 TTCTCGCCACGCTTCGCCGGCTTTCCCGTCAAGCTCTAAAT 112

RESULT 14
LOCUS      AQ280347
DEFINITION CITBI-E1-2518K3.TR CITBI-E1 Homo sapiens genomic clone 2518K3,
            genomic survey sequence.
ACCESSION  AQ280347
VERSION    AQ280347.1 GI:3906166
KEYWORDS   GSS.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 211)

```

```

68 TTCTCGCCACGCTTCGCCGGCTTTCCCGTCAAGCTCTAAAT 108

RESULT 13
LOCUS      AQ278442
DEFINITION CITBI-E1-2519N1.TR CITBI-E1 Homo sapiens genomic clone 2519N1,
            genomic survey sequence.
ACCESSION  AQ278442
VERSION    AQ278442.1 GI:3904410
KEYWORDS   GSS.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 200)
AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
            Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
            Venter,J.C.
TITLE     Use of a random human BAC End Sequence Database for Sequence-Ready
            Map Building
JOURNAL    Unpublished (1998)
COMMENT    Contact: Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: mdadams@tigr.org
            Clones are available from Research Genetics (info@resgen.com). BAC
            end search page:
            http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
            Seq primer: M13 Reverse
            Class: BAC ends.

FEATURES             Location/Qualifiers
     source            1..200
     map_feature        1..200
     other_GSSs        CITBI-E1-2519N1.TF
     Contact: Mark Adams
     Department of Eukaryotic Genomics
     The Institute for Genomic Research
     9712 Medical Center Dr., Rockville, MD 20850, USA
     Tel: 301 838 0200
     Fax: 301 838 0208
     Email: mdadams@tigr.org
     Clones are available from Research Genetics (info@resgen.com). BAC
     end search page:
     http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
     Seq primer: M13 Reverse
     Class: BAC ends.

FEATURES             Location/Qualifiers
     source            1..200
     map_feature        1..200
     other_GSSs        CITBI-E1-2519N1.TF
     Contact: Mark Adams
     Department of Eukaryotic Genomics
     The Institute for Genomic Research
     9712 Medical Center Dr., Rockville, MD 20850, USA
     Tel: 301 838 0200
     Fax: 301 838 0208
     Email: mdadams@tigr.org
     Clones are available from Research Genetics (info@resgen.com). BAC
     end search page:
     http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
     Seq primer: M13 Reverse
     Class: BAC ends.

ORIGIN
Query Match      100.0%; Score 101; DB 8; Length 200;
Best Local Similarity 100.0%; Pred. No. 9.5e-18;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTCCTTTTCGCTTTCTTCCCTTCCT 60
Db 12 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCCGCTCCTTTTCGCTTTCTTCCCTTCCT 71

Qy 61 TTCTCGCCACGCTTCGCCGGCTTTCCCGTCAAGCTCTAAAT 101
Db 72 TTCTCGCCACGCTTCGCCGGCTTTCCCGTCAAGCTCTAAAT 112

RESULT 14
LOCUS      AQ280347
DEFINITION CITBI-E1-2518K3.TR CITBI-E1 Homo sapiens genomic clone 2518K3,
            genomic survey sequence.
ACCESSION  AQ280347
VERSION    AQ280347.1 GI:3906166
KEYWORDS   GSS.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 211)

```

AUTHORS Adams, M.D., Rounsley, S.D., Zhao, S., Bass, S., Linher, K., Golden, K., Berry, K., Granger, D., Suh, E., Wible, C., Shizuya, H., Simon, M. and Venter, J.C.

TITLE Use of a random human BAC End Sequence Database for Sequence-Ready Map Building

JOURNAL COMMENT Unpublished (1998)
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: mdamas@tigr.org

Clones are available from Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tcdb/hungen/bac_end_search/bac_end_search.html.

Seq primer: M13 Reverse
 Class: BAC ends.

FEATURES

source
 1. .211
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="2518K3"
 /sex="male"
 /cell_type="sperm"
 /clone_lib="CITBI-E1"
 /note="Vector: pBelobAC11; Site_1: EcoRI; Site_2: EcoRI; Caltech Human BAC Library D"

ORIGIN

Query Match 100.0%; Score 101; DB 8; Length 211;
 Best Local Similarity 100.0%; Pred. No. 9.6e-18;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCCTTTCGCTTCTCCCTTCCT 60
 |||||
 Db 8 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCCTTTCGCTTCTCCCTTCCT 67
 |||||

Qy 61 TTCTCGCCACGTTCCCGCGCTTCCCGCTCAAGCTCTAAAT 101
 |||||
 Db 68 TTCTCGCCACGTTCCCGCGCTTCCCGCTCAAGCTCTAAAT 108
 |||||

RESULT 15
 BI677411

LOCUS BI677411 217 bp mRNA linear EST 17-SEP-2001

DEFINITION ie30903.y1 Kaestner ngn3 wt Mus musculus cDNA 5', mRNA sequence.

ACCESSION BI677411

VERSION BI677411.1 GI:15630318

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 217)
 Melton, D., Brown, J., Kenty, G., Brestelli, J., Gradwohl, G., Clifton, S., Lenishka, I., Searce, M., Pape, D., Wylie, T., Martin, J., Blistain, A., Hillier, L., Marra, M., Rittner, E., Rittner, E., Rittner, E., Bennett, J., Schmitt, A., Theising, B., Rittner, E., Rittner, E., Rittner, J., Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagareishvili, R., Williams, T., Jackson, Y. and Bowers, Y.
 Endocrine Pancreas Consortium
 Unpublished (2000)
 Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
 Endocrine Pancreas Consortium
 Harvard University, Howard Hughes Medical Institute
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138
 Tel: 617-495-1812
 Fax: 617-495-8557
 Email: dmelton@biohpc.harvard.edu
 Pancreas was obtained from Gerard Gradwohl (PNAS 97 P1607-1611.

2000) Library was constructed by Catherine Lee DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Dr. Marie Searce (msearce@mail.med.upenn.edu)
 Seq primer: -40RP from Gibco
 High quality sequence stop: 207.

FEATURES

Location/Qualifiers
 1. .217
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="129/Sv x CD1"
 /db_xref="taxon:10090"
 /dev_stage="p.c. 14.5"
 /lab_host="E. coli-DH12S (GIBCO)"
 /clone_lib="Kaestner ngn3 wt"
 /note="Organ: pancreas; Vector: pSPORT1 (GIBCO); Site_1: Not I; Site_2: Sal I; The library was prepared by Catherine S. Lee and has not been published. The pancreas was obtained from Gerard Gradwohl (PNAS 97 P1607-1611, 2000). The cDNA's were prepared with an oligo containing a NotI site, and SalI linkers were added to the ends. The inserts were cut with NotI before being cloned into the NotI-SalI sites in the vectors. This is one of two libraries, ngn3 wt and ngn3 -/- . The wt library is in pSPORT1, T7 promoter is 5'."

ORIGIN

Query Match 100.0%; Score 101; DB 4; Length 217;
 Best Local Similarity 100.0%; Pred. No. 9.6e-18;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCCTTTCGCTTCTCCCTTCCT 60
 |||||
 Db 28 GCGTGACCGCTACACTTGCAGCGCCCTAGCGCGCTCCTTTCGCTTCTCCCTTCCT 87
 |||||

Qy 61 TTCTCGCCACGTTCCCGCGCTTCCCGCTCAAGCTCTAAAT 101
 |||||
 Db 88 TTCTCGCCACGTTCCCGCGCTTCCCGCTCAAGCTCTAAAT 128
 |||||

Search completed: July 14, 2005, 23:23:19
 Job time : 966.667 secs

THIS PAGE BLANK (USPTO)


```
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 142)
AUTHORS Kunsch,C.A., Choi,G.A., Barash,S.C., Dillon,P.J., Fannon,M.R. and Rosen,C.A.
TITLE Staphylococcus aureus polynucleotides and sequences
JOURNAL Patent: US 6737248-A 2608 18-MAY-2004;
FEATURES Location/Qualifiers
    source 1..142
            /organism="unknown"
            /mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 142;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAATAG 60
    |||||
Db 107 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAATAG 48
    |||||

Qy 61 GGGTTCGGCGCACATTTCCCGGAAAAGTGCCACCTGACGTC 101
    |||||
Db 47 GGGTTCGGCGCACATTTCCCGGAAAAGTGCCACCTGACGTC 7
    |||||

RESULT 3
E00019/c
LOCUS DNA coding for Escherichia coli penicillinase.
DEFINITION E00019
ACCESSION E00019
VERSION E00019.1 GI:2168327
KEYWORDS JP 1981154999-A/2.
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE 1 (bases 1 to 228)
AUTHORS Uorutaa,G. and Karen,T.
TITLE SYNTHESIS OF SELECTIVE AGED PROTEIN OR POLYPEPTIDE IN BACTERIA
JOURNAL Patent: JP 1981154999-A 2 30-NOV-1981;
UNIV HARVARD
COMMENT OS Escherichia coli
PN JP 1981154999-A/2
PD 30-NOV-1981
PF 09-APR-1981 JP 1981052488
PI 11-APR-1980 US 80 139225
PC UORUTAA GIRUBAATO, KAREN TARUMATSUJI
PC C12P21/00,C07H21/00,C12N1/00,C12N15/00//C12R1/19; CC
strandedness: Double;
CC topology: Linear;
CC anti-sense: No;
CC *source: clone=pKT218;
FH Key Location/Qualifiers
FH CDS 210..>228
FT /product='E.coli penicillinase'.
FEATURES
    source 1..228
            Location/Qualifiers
            1..228
                /organism="Escherichia coli"
                /mol_type="genomic DNA"
                /db_xref="taxon:562"

ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 228;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAATAG 60
    |||||
Db 175 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAATAG 116
    |||||

Qy 61 GGGTTCGGCGCACATTTCCCGGAAAAGTGCCACCTGACGTC 101
    |||||
```

```
Db 115 GGGTTCGGCGCACATTTCCCGGAAAAGTGCCACCTGACGTC 75
    |||||

RESULT 4
PMOENDO/c
LOCUS DNA 240 bp linear BCT 26-APR-1993
DEFINITION Plasmid pMM110 region of endo VII cleavage sites near cruciform structures.
ACCESSION M10199
VERSION M10199.1 GI:150826
KEYWORDS
SOURCE Plasmid pMM110
ORGANISM Plasmid pMM110
other sequences: plasmids.
REFERENCE 1 (bases 1 to 240)
AUTHORS Kemper,B., Jensch,F., von Depka-Prondzynski,M., Fritz,H.J., Borgmeyer,U. and Mizuuchi,K.
TITLE Resolution of Holliday structures by endonuclease VII as observed in interactions with cruciform DNA
JOURNAL Cold Spring Harb. Symp. Quant. Biol. 49, 815-825 (1984)
MEDLINE 85153063
PUBMED 6397324
COMMENT Original source text: Plasmid pMM110 DNA.
FEATURES
    source 1..240
            Location/Qualifiers
            1..240
                /organism="Plasmid pMM110"
                /mol_type="genomic DNA"
                /db_xref="taxon:2599"
                /plasmid="Plasmid pMM110"

ORIGIN Unreported.

Query Match 100.0%; Score 101; DB 1; Length 240;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTTCGAATGTTATTAGAAAAATAAACAATAG 60
    |||||
Db 151 AGGGTTATTGCTCATGAGCGGATACATATTTCGAATGTTATTAGAAAAATAAACAATAG 92
    |||||

Qy 61 GGGTTCGGCGCACATTTCCCGGAAAAGTGCCACCTGACGTC 101
    |||||
Db 91 GGGTTCGGCGCACATTTCCCGGAAAAGTGCCACCTGACGTC 51
    |||||

RESULT 5
E00018/c
LOCUS DNA coding for Escherichia coli penicillinase.
DEFINITION E00018
ACCESSION E00018
VERSION E00018.1 GI:2168326
KEYWORDS JP 1981154999-A/1.
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE 1 (bases 1 to 251)
AUTHORS Uorutaa,G. and Karen,T.
TITLE SYNTHESIS OF SELECTIVE AGED PROTEIN OR POLYPEPTIDE IN BACTERIA
JOURNAL Patent: JP 1981154999-A 1 30-NOV-1981;
UNIV HARVARD
COMMENT OS Escherichia coli
PN JP 1981154999-A/1
PD 30-NOV-1981
PF 09-APR-1981 JP 1981052488
PI 11-APR-1980 US 80 139225
PC UORUTAA GIRUBAATO, KAREN TARUMATSUJI
PC C12P21/00,C07H21/00,C12N1/00,C12N15/00//C12R1/19; CC
strandedness: Double;
CC topology: Linear;
CC anti-sense: No;
CC fragment type: N-Terminal Fragment;
CC *source: clone=pKT241;
```

```

FH Key Location/Qualifiers
FH CDS 210..>252
FT /product='E.coli penicillinase' FT
TATA_signal 190..196
FEATURES             Location/Qualifiers
     source          1..251
                     /organism="Escherichia coli"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:562"
ORIGIN
Query Match          100.0%; Score 101; DB 6; Length 251;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTAGAAAATAAACAATAG 60
    |||||||
Db 175 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTAGAAAATAAACAATAG 116
    |||||||
Qy 61 GGGTTCGCGCACATTTCCCGGAAAAGTGCACCTGACGTC 101
    |||||||
Db 115 GGGTTCGCGCACATTTCCCGGAAAAGTGCACCTGACGTC 75

RESULT 6
I01644/c
LOCUS                251 bp ss-DNA linear PAT 18-MAY-1993
DEFINITION           Sequence 1 from Patent US 4338397.
ACCESSION            I01644
VERSION              I01644.1 GI:267685
KEYWORDS
SOURCE               Unknown.
ORGANISM              Unclassified.
REFERENCE             1 (bases 1 to 251)
AUTHORS              Gilbert,W. and Talmadge,K.
TITLE                Mature protein synthesis
JOURNAL              Patent: US 4338397-A 1 06-JUL-1982;
                    President and Fellows of Harvard College; Cambridge, MA
FEATURES             Location/Qualifiers
     source          1..251
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN
Query Match          100.0%; Score 101; DB 6; Length 251;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTAGAAAATAAACAATAG 60
    |||||||
Db 175 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTAGAAAATAAACAATAG 116
    |||||||
Qy 61 GGGTTCGCGCACATTTCCCGGAAAAGTGCACCTGACGTC 101
    |||||||
Db 115 GGGTTCGCGCACATTTCCCGGAAAAGTGCACCTGACGTC 75

RESULT 7
HUMUT5345
LOCUS                344 bp DNA linear STS 26-JUL-1993
DEFINITION           Human chromosome 8 STS UT5345, sequence tagged site.
ACCESSION            L18624
VERSION              L18624.1 GI:308338
KEYWORDS              STS; PCR primer; STS sequence; microsatellite marker;
                    microsatellite repeat; repeat polymorphism; sequence tagged site.
SOURCE               Homo sapiens
ORGANISM              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE             1 (bases 1 to 344)
AUTHORS              Gerken,S.C., Matsunami,N., Lawrence,E., Carlson,M., Moore,M.,

```

```

Ballard,L., Melis,R., Robertson,M., Bradley,P., Elsner,T.,
Tingey,A., Rodriguez,P., Albertsen,H., Lalouel,J.-M. and White,R.
Genetic and physical mapping of simple sequence repeat containing
sequence tagged sites from the human genome
Unpublished (1993)
Original source text: Homo sapiens DNA.
Submitted by: Utah Center for Human Genome Research University of
Utah, Dept. of Human Genetics
2160 Eccles Institute of Human Genetics
Salt Lake City, UT 84112
e-mail: sts@corona.med.utah.edu
Primer A: GAGCAAAAACAGGAGGCAATATGC
Primer B: TTCGGGAAATGTGCCGGAACC
32P-label: B Primer
PCR Profile:
Initial Denaturation: 94C 300sec
PCR Cycles: 30
Denaturation: 94C 10sec
Annealing: 60C 10sec
Extension: 72C 20sec
Mg++: 2mM
Gel: Acrylamide 7%, Formamide 32%, Urea 34%
Alleles: 2
FEATURES             Location/Qualifiers
     source          1..344
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /map="8"
                     /standard_name="STS UT5345"
                     /complement(202..224)
     primer_bind     36..224
     primer_bind     36..60
ORIGIN
Query Match          100.0%; Score 101; DB 11; Length 344;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTAGAAAATAAACAATAG 60
    |||||||
Db 141 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTATTAGAAAATAAACAATAG 200
    |||||||
Qy 61 GGGTTCGCGCACATTTCCCGGAAAAGTGCACCTGACGTC 101
    |||||||
Db 201 GGGTTCGCGCACATTTCCCGGAAAAGTGCACCTGACGTC 241

RESULT 8
BD195256/c
LOCUS                400 bp DNA linear PAT 17-JUL-2003
DEFINITION           Nucleotide sequence of Escherichia coli pathogenicity islands.
ACCESSION            BD195256
VERSION              BD195256.1 GI:33005021
KEYWORDS              JP 2002513277-A/43.
SOURCE               unidentified
ORGANISM              unidentified.
REFERENCE             1 (bases 1 to 400)
AUTHORS              Dillon,J., Choi,G.H. and Welch,R.A.
TITLE                Nucleotide sequence of Escherichia coli pathogenicity islands
JOURNAL              Patent: JP 2002513277-A 43 08-MAY-2002;
                    HUMAN GENOME SCIENCES INC,WISCONSIN ALUMNI RESEARCH FOUNDATION
COMMENT              OS Unidentified
                    PN JP 2002513277-A/43
                    PD 08-MAY-2002
                    PR 21-NOV-1997 JP 1998523916
                    PI 60/031626,14-OCT-1997 US 60/061953
                    PC CL2N15/11.C12N15/63,C07K16/12.G01N33/569,G06F17/30.G11B7/00
                    CC Strandedness: Double;
                    CC Topology: Linear;
                    CC Nucleotide sequence of Escherichia coli pathogenicity islands

```

```

FH Key Location/Qualifiers
FT source 1..400
FT Location/Qualifiers
FEATURES
source 1..400
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 400;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGTTATTGCTCATGAGCGGATACATATTGTAATGATTTAGAAAAATAAACAATAG 60
Db 165 AGGTTATTGCTCATGAGCGGATACATATTGTAATGATTTAGAAAAATAAACAATAG 106

Qy 61 GGGTTCCGCGACATTTCCCGGAAAGTGCACCTGACGTC 101
Db 105 GGGTTCCGCGACATTTCCCGGAAAGTGCACCTGACGTC 65

RESULT 9
E00892/c
LOCUS E00892 456 bp DNA linear PAT 29-SEP-1997
DEFINITION Synthetic DNA encoding fused polypeptide between E coli
beta-lactamase and human beta-urogastrone.
ACCESSION E00892
VERSION E00892.1 GI:2169153
KEYWORDS JP 1986149089-A/1.
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 456)
AUTHORS Okai,H., Momota,Y., Kumakura,T., Tochifusa,N., Kitazawa,T.,
Ojida,K. and Matsushiro,A.
TITLE POLYPEPTIDE SECRETION DEVELOPMENT VECTOR, MICROORGANISM TRANSFORMED
WITH SAME, AND PRODUCTION OF POLYPEPTIDE WITH MICROORGANISM
JOURNAL Patent: JP 1986149089-A 1 07-JUL-1986;
EARTH CHEM CORP LTD
COMMENT OS Artificial gene
OC Artificial sequence; Genes.
PN JP 1986149089-A/1
PD 07-JUL-1986
PF 21-DEC-1984 JP 1984271206
PI OKAI HIDEO, MOMOTA YUTAKA, KUMAKURA TAKESHI,
PI TOCHIFUSA NORIYUKI,
PI KITAZAWA TOSHIKI, OJIDA KAZUHIRO, MATSUSHIRO AIZO PC
C12N15/00,C12N1/20,C12P21/00,(C12N1/20,C12R1:19),(C12P21/00,PC
C12R1:19);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: strain=HB101;
CC *source: clones=pVG201;
CC Feature is identified by experimental;
FH Key Location/Qualifiers
FH promoter 125..170
FT of beta-lactamase
FT RBS 200..203
FT CDS 209..438
FT /product='beta-urogastrone precursor' FT
FT sig_peptide 209..277
FT /product='signal peptide of beta-lactonase' FT
FT mat_peptide 278..435
FT /product='beta-urogastrone mature peptide'.
FT Location/Qualifiers
FT 1..456
/organism="synthetic construct"
/mol_type="genomic DNA"

FEATURES
source

```

```

ORIGIN
/db_xref="taxon:32630"

Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGTTATTGCTCATGAGCGGATACATATTGTAATGATTTAGAAAAATAAACAATAG 60
Db 173 AGGTTATTGCTCATGAGCGGATACATATTGTAATGATTTAGAAAAATAAACAATAG 114

Qy 61 GGGTTCCGCGACATTTCCCGGAAAGTGCACCTGACGTC 101
Db 113 GGGTTCCGCGACATTTCCCGGAAAGTGCACCTGACGTC 73

RESULT 10
E01156/c
LOCUS E01156 456 bp DNA linear PAT 29-SEP-1997
DEFINITION DNA fragment which secretes beta urogastrone.
ACCESSION E01156
VERSION E01156.1 GI:2169415
KEYWORDS JP 1987083890-A/1.
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 456)
AUTHORS Yoshikawa,K., Momota,Y., Kajifusa,N., Koide,T. and Okai,H.
TITLE POLYPEPTIDE SECRETION EXPRESSION VECTOR, MICROORGANISM TRANSFORMED
BY SAID VECTOR AND PRODUCTION OF POLYPEPTIDE USING SAID
JOURNAL Patent: JP 1987083890-A 1 17-APR-1987;
EARTH CHEM CORP LTD
COMMENT OS Artificial gene
OC Artificial sequence; Genes.
PN JP 1987083890-A/1
PD 17-APR-1987
PF 09-OCT-1985 JP 1985225393
PI YOSHIKAWA KAZUTOSHI, MOMOTA YUTAKA, KAJIFUSA NORIYUKI, PI
KOIDE TAKAO,
PI OKAI HIDEO
PC C12N15/00,C12N1/20,C12P21/00,(C12N1/20,C12R1:19),(C12N1/20,PC
C12R1:125);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: clone=pUG201;
FH Key Location/Qualifiers
FH promoter 125..170
FT /note='beta lactamase promoter' FT RBS
FT CDS 200..204
FT /product='beta urogastrone'
FT sig_peptide 209..277
FT mat_peptide 278..436
FT /product='beta urogastrone'.
FEATURES
source
Location/Qualifiers
1..456
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGTTATTGCTCATGAGCGGATACATATTGTAATGATTTAGAAAAATAAACAATAG 60
Db 173 AGGTTATTGCTCATGAGCGGATACATATTGTAATGATTTAGAAAAATAAACAATAG 114

Qy 61 GGGTTCCGCGACATTTCCCGGAAAGTGCACCTGACGTC 101

```


Db 113 GGGTTCCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 73
|||||
E01274 456 bp DNA linear PAT 29-SEP-1997
DEFINITION DNA encoding beta-urogastron fused with DNA encoding a promoter and
signal peptide of beta-lactamase.
E01274
ACCESSION E01274.1 GI:2169533
VERSION JP 1987179398-A/1.
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 456)
AUTHORS Okai, H., Kumakura, T., Kawamoto, S., Adachi, S., Matsubara, A.,
Ojida, K., Yano, M., Mihara, S., Matsuehiro, A. and Yanaihara, N.
TITLE PRODUCTION OF BETA-UROGASTRONE
JOURNAL PATENT: JP 1987179398-A 1 06-AUG-1987;
COMMENT EARTH CHEM CORP LTD
OS Artificial gene
OC Artificial sequence; Genes.
OS Homo sapiens
PN JP 1987179398-A/1
PD 06-AUG-1987
PF 31-JAN-1986 JP 1986021032
PI OKAI HIDEO, KUMAKURA TAKESHI, KAWAMOTO SHOICHI, ADACHI SHOJI,
MATSUBARA AKIMASA, OJIDA KAZUHIDE, YANO MAKI, MIHARA SHIGERU,
MATSUSHIRO AIZO, YANAIHARA NOBORU
PC C12P21/00, C12N15/00, (C12P21/00, C12R1:91);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No; Location/Qualifiers
FH Key
FH
FT Promoter 125..170
FT RBS 200..203
FT sig_peptide 209..277
FT mat_peptide 278..436
FT CDS 209..439
FT /product='beta-urogastron'.
FT /product='beta-urogastron'.
FEATURES Location/Qualifiers
source 1..456
/organism='synthetic construct'
/mol_type='genomic DNA'
/db_xref='taxon:32630'
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20; Mismatches 0; Indels 0; Gaps 0;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGATTAGAAAATAAACAATAG 60
|||||
Db 173 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGATTAGAAAATAAACAATAG 114
|||||
Qy 61 GGGTTCCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101
|||||
Db 113 GGGTTCCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 73
|||||
RESULT 13
AX260098/c
LOCUS AX260098 466 bp DNA linear PAT 26-OCT-2001
DEFINITION Sequence 60 from Patent WO0172774.
ACCESSION AX260098
VERSION AX260098.1 GI:16509129
KEYWORDS
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1
AUTHORS Deak, P., Glover, D.M. and Midgley, C.
TITLE Cell cycle progression proteins
JOURNAL Patent: WO 0172774-A 60 04-OCT-2001;
CYCLACEL Limited (GB)
FEATURES Location/Qualifiers
source 1..466
/organism='Drosophila melanogaster'
/mol_type='unassigned DNA'

Db 113 GGGTTCCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 73
|||||
E01274 456 bp DNA linear PAT 29-SEP-1997
DEFINITION DNA encoding beta-urogastron fused with DNA encoding a promoter and
signal peptide of beta-lactamase.
E01274
ACCESSION E01274.1 GI:2169533
VERSION JP 1987179398-A/1.
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 456)
AUTHORS Okai, H., Kumakura, T., Kawamoto, S., Adachi, S., Matsubara, A.,
Ojida, K., Yano, M., Mihara, S., Matsuehiro, A. and Yanaihara, N.
TITLE PRODUCTION OF BETA-UROGASTRONE
JOURNAL PATENT: JP 1987179398-A 1 06-AUG-1987;
COMMENT EARTH CHEM CORP LTD
OS Artificial gene
OC Artificial sequence; Genes.
OS Homo sapiens
PN JP 1987179398-A/1
PD 06-AUG-1987
PF 31-JAN-1986 JP 1986021032
PI OKAI HIDEO, KUMAKURA TAKESHI, KAWAMOTO SHOICHI, ADACHI SHOJI,
MATSUBARA AKIMASA, OJIDA KAZUHIDE, YANO MAKI, MIHARA SHIGERU,
MATSUSHIRO AIZO, YANAIHARA NOBORU
PC C12P21/00, C12N15/00, (C12P21/00, C12R1:91);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No; Location/Qualifiers
FH Key
FH
FT Promoter 125..170
FT RBS 200..203
FT sig_peptide 209..277
FT mat_peptide 278..436
FT CDS 209..439
FT /product='beta-urogastron'.
FT /product='beta-urogastron'.
FEATURES Location/Qualifiers
source 1..456
/organism='synthetic construct'
/mol_type='genomic DNA'
/db_xref='taxon:32630'
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20; Mismatches 0; Indels 0; Gaps 0;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGATTAGAAAATAAACAATAG 60
|||||
Db 173 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGATTAGAAAATAAACAATAG 114
|||||
Qy 61 GGGTTCCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101
|||||
Db 113 GGGTTCCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 73
|||||
RESULT 12
E01302/c
LOCUS E01302 456 bp DNA linear PAT 29-SEP-1997
DEFINITION DNA encoding human beta-urogastrone fused with DNA encoding
promoter and signal peptide of beta-lactamase.
E01302
ACCESSION E01302.1 GI:2169561
VERSION JP 1987190083-A/1.
KEYWORDS synthetic construct
SOURCE synthetic construct

```
ORIGIN
/db_xref="taxon:7227"

Query Match      100.0%; Score 101; DB 6; Length 466;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 60
    |||||
Db 280 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 221
    |||||

Qy 61 GGGTTCGGCGCACATTTCCCGAAAAAGTCCACCTGACGTC 101
    |||||
Db 220 GGGTTCGGCGCACATTTCCCGAAAAAGTCCACCTGACGTC 180
    |||||

RESULT 14
AX260150/c      AX260150      573 bp      DNA      linear      PAT 26-OCT-2001
LOCUS
DEFINITION      Sequence 112 from Patent WO0172774.
ACCESSION      AX260150
VERSION      AX260150.1 GI:16509172
KEYWORDS
SOURCE
ORGANISM
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1
AUTHORS      Deak, P., Glover, D.M. and Midgley, C.
TITLE      Cell cycle progression proteins
JOURNAL      Patent: WO 0172774-A 112 04-OCT-2001;
Cyclacel Limited (GB)
FEATURES
source
1..573
/organism="Drosophila melanogaster"
/mol_type="unassigned DNA"
/db_xref="taxon:7227"

ORIGIN

Query Match      100.0%; Score 101; DB 6; Length 573;
Best Local Similarity 100.0%; Pred. No. 8.8e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 60
    |||||
Db 355 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 296
    |||||

Qy 61 GGGTTCGGCGCACATTTCCCGAAAAAGTCCACCTGACGTC 101
    |||||
Db 295 GGGTTCGGCGCACATTTCCCGAAAAAGTCCACCTGACGTC 255
    |||||

RESULT 15
A43586
LOCUS
DEFINITION      Sequence 11 from Patent WO9507357.
ACCESSION      A43586
VERSION      A43586.1 GI:2298779
KEYWORDS
SOURCE
ORGANISM
Cuphea lanceolata
Cuphea lanceolata
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; Myrtales; Lythraceae; Cuphea.
1 (bases 1 to 693)
Toepfer, R., Bautor, J., Bothmann, H., Filsak, E.,
Hoerlcke-Grandpierre, C., Klein, B., Martini, N., Mueller, A.,
Schulte, W., Voetz, M., Walek, J. and Schell, J.
PROMOTERS
TITLE      Patent: WO 9507357-A 11 16-MAR-1995;
JOURNAL      MAX PLANCK GESELLSCHAFT (DB)
COMMENT      Other publication CA 2169093 950316
```

```
FEATURES
source
Other publication AU 7615494 950327.
Location/Qualifiers
1..693
/organism="Cuphea lanceolata"
/mol_type="unassigned DNA"
/db_xref="taxon:3930"
/clone="CLKASIG8"
/clone_lib="Genomic Lambda Fix II"

ORIGIN.

Query Match      100.0%; Score 101; DB 6; Length 693;
Best Local Similarity 100.0%; Pred. No. 8.8e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 60
    |||||
Db 592 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 651
    |||||

Qy 61 GGGTTCGGCGCACATTTCCCGAAAAAGTCCACCTGACGTC 101
    |||||
Db 652 GGGTTCGGCGCACATTTCCCGAAAAAGTCCACCTGACGTC 692
    |||||

Search completed: July 14, 2005, 14:03:36
Job time : 756.618 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:35:42 ; Search time 142.398 Seconds
(without alignments)
4198.742 Million cell updates/sec

Title: US-09-482-682-64_COPY_7131_7231
Perfect score: 101
Sequence: 1 aggtttattgtctcatgacg.....gaaagtgcacgtgacgtc 101

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002as:*
- 7: geneseqn2002bs:*
- 8: geneseqn2003as:*
- 9: geneseqn2003bs:*
- 10: geneseqn2003cs:*
- 11: geneseqn2003ds:*
- 12: geneseqn2004as:*
- 13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	101	100.0	142	2 AAV76919	AAV76919 Staphyloc
C 2	101	100.0	228	1 AAN10032	Aan10032 Sequence
C 3	101	100.0	251	1 AAN10031	Aan10031 Sequence
C 4	101	100.0	400	2 AAV31229	AAV31229 E. coli J
C 5	101	100.0	456	1 AAN60624	Aan60624 Plasmid p
C 6	101	100.0	456	1 AAN71080	Aan71080 Sequence
C 7	101	100.0	456	1 AAN70833	Aan70833 Beta-urog
C 8	101	100.0	456	1 AAN81765	Aan81765 Sequence
C 9	101	100.0	466	6 ABA90413	ABa90413 Drosophil
C 10	101	100.0	487	2 AAX21173	Aax21173 Polynucle
C 11	101	100.0	535	2 AAX21149	Aax21149 Polynucle
C 12	101	100.0	573	6 ABA90456	ABa90456 Drosophil
C 13	101	100.0	605	12 ADH58311	ADH58311 Electroph
C 14	101	100.0	776	4 AAS30560	Aas30560 DNA encod
C 15	101	100.0	776	4 AAS27819	Aas27819 DNA encod
C 16	101	100.0	776	4 ABK42984	ABk42984 Genomic s
C 17	101	100.0	776	4 AAL07344	Aal07344 Human rep
C 18	101	100.0	776	4 AAL03229	Aal03229 Human rep
C 19	101	100.0	776	4 AAL06588	Aal06588 Human rep
C 20	101	100.0	776	4 AAL07340	Aal07340 Human rep

C 21	101	100.0	776	5 ABA14573	ABa14573 Human ner
C 22	101	100.0	776	5 AAS34681	Aas34681 Human DNA
C 23	101	100.0	776	8 ADA41574	Ada41574 Human sec
C 24	101	100.0	776	8 ACC50905	Acc50905 Human sec
C 25	101	100.0	776	8 ABZ71508	Abz71508 Secreted
C 26	101	100.0	776	9 ADB91869	ADB91869 Human sec
C 27	101	100.0	776	9 ADB61140	ADB61140 Connectiv
C 28	101	100.0	776	10 ADB94622	ADB94622 Novel hum
C 29	101	100.0	776	10 ADC74663	Adc74663 Human sec
C 30	101	100.0	776	10 ADA57709	Ada57709 BAC fragm
C 31	101	100.0	776	12 ADN41551	Adn41551 Novel hum
C 32	101	100.0	845	4 AAS30559	Aas30559 DNA encod
C 33	101	100.0	845	4 AAS27818	Aas27818 DNA encod
C 34	101	100.0	845	4 ABK42983	ABk42983 Genomic s
C 35	101	100.0	845	4 AAS41807	Aas41807 Genomic s
C 36	101	100.0	845	4 AAS41855	Aas41855 Genomic s
C 37	101	100.0	845	4 AAK85485	Aak85485 Human imm
C 38	101	100.0	845	4 AAK85434	Aak85434 Human imm
C 39	101	100.0	845	4 AAL07343	Aal07343 Human rep
C 40	101	100.0	845	4 AAL06587	Aal06587 Human rep
C 41	101	100.0	845	4 AAL07339	Aal07339 Human rep
C 42	101	100.0	845	4 AAL03228	Aal03228 Human rep
C 43	101	100.0	845	5 ABA14572	ABa14572 Human ner
C 44	101	100.0	845	5 AAS34680	Aas34680 Human DNA
C 45	101	100.0	845	9 ADB61139	ADB61139 Connectiv

ALIGNMENTS

RESULT 1
AAV76919/c
ID AAV76919 standard; DNA; 142 BP.

XX AC AAV76919;
XX DT 16-MAR-1999 (first entry)
XX DE Staphylococcus aureus contig SEQ ID #2608.
XX KW Computer readable medium; vaccine; S.aureus infection; immunodetection;
KW skin infection; eyelid infection; food poisoning; osteomyelitis; therapy;
KW toxic shock syndrome; ds.
XX OS Staphylococcus aureus.
XX PN EP786519-A2.
XX PD 30-JUL-1997.
XX PF 07-JAN-1997; 97EP-00100117.
XX PR 05-JAN-1996; 96US-0009861P.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Kunsch CA, Choi GH, Barash SC, Dillon PJ, Fannon MR, Rosen CA;
XX DR WPI; 1997-374922/35.
XX PT Polynucleotide(s) and proteins derived from Staphylococcus aureus -
XX stored on computer readable medium and used in the production of anti-
XX S.aureus vaccines.
XX PS Claim 1; Page 2287; 3271pp; English.
XX CC This sequence represents one of 5191 Staphylococcus aureus DNA sequences
CC of the invention. The DNA sequences are recorded on a computer readable
CC medium, preferably selected from a floppy or hard disk, random access
CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
CC the S.aureus DNA sequences allows putative functions to be assigned so
CC that protein-encoding or regulatory regions of commercial, therapeutic or

CC industrial importance can be obtained. Specifically, sequences which are
 CC likely to encode antigens have been identified and these polypeptides can
 CC be used in a vaccine composition against *S. aureus* infection. The
 CC polypeptides can also be used in a kit for the immunodetection of
 CC *S. aureus* in a sample. *S. aureus* is implicated in numerous human diseases,
 CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,
 CC skin and surgical wound infections, scalded skin syndrome, toxic shock
 CC syndrome, etc. Organisms transformed with the DNA sequences can be used
 CC for recombinant production of the polypeptides. The new DNA sequences
 CC (and their fragments) are useful as primers or probes for isolating
 CC homologues of any of the *S. aureus* DNA sequences contained on the computer
 CC readable medium

SQ Sequence 142 BP; 45 A; 25 C; 26 G; 45 T; 0 U; 1 Other;

Query Match 100.0%; Score 101; DB 2; Length 142;
 Best Local Similarity 100.0%; Pred. No. 2.1e-21;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGCTATTAGAAAAATAACAATAG 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 107 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGCTATTAGAAAAATAACAATAG 48
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 61 GGGTTCCGGCGACATTTCCCGAAAAAGTGCACCTGACGTC 101
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 47 GGGTTCCGGCGACATTTCCCGAAAAAGTGCACCTGACGTC 7
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 2
 AAN10032/C
 ID AAN10032 standard; DNA; 228 BP.
 XX AC AAN10032;
 XX DT 13-AUG-1992 (first entry)
 XX DE Sequence of the pKT218 EcoRI-PstI penicillinase gene fragment.
 XX KW Cloning vehicle; bacterial vector; transformed host; penicillinase;
 XX KW insulin; ds.
 XX OS *Escherichia coli*.
 XX FH Key Location/Qualifiers
 FT misc_feature 1..4
 FT /*tag= a
 FT /label= sticky end
 FT misc_feature 225..228
 FT /*tag= b
 FT /label= sticky end
 XX EP38182-A.
 XX PN 21-OCT-1981.
 XX PD 09-APR-1981; 81EP-00301561.
 XX PF 11-APR-1980; 80US-00139225.
 XX PR (HARD) HARVARD COLLEGE.
 XX PA Gilbert W, Talmadge K;
 XX PI WPI; 1981-80125D/44.
 XX DR P-PSDB; AAP10039.
 XX PS Synthesis of mature protein or polypeptide - by using bacterial host
 PT transformed by cloned vehicle contg. DNA fragment etc.
 XX Example; Fig 3; 34pp; English.
 XX CC The closest identifiable promoter for the penicillinase gene in pKT241
 CC (AAN10031) is located in the region 14 to 20 nucleotides before its

CC translational start signal. In the examples, the 3' end of pKT241 was
 CC attached to the signal DNA sequence of the DNA fragment (19) for rat
 CC preproinsulin (see AAN10033). The closest identifiable promoter for the
 CC penicillinase gene in pKT218 (AAN10032) is located in the region 14 to 20
 CC nucleotides before its translational start signal. In the examples, the
 CC 3' end of pKT218 was attached to the signal DNA sequence of the DNA
 CC fragment (CB6) for rat preproinsulin (see AAN10034)

SQ Sequence 228 BP; 72 A; 37 C; 46 G; 73 T; 0 U; 0 Other;
 Query Match 100.0%; Score 101; DB 1; Length 228;
 Best Local Similarity 100.0%; Pred. No. 2.3e-21;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGCTATTAGAAAAATAACAATAG 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 175 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGCTATTAGAAAAATAACAATAG 116
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 61 GGGTTCCGGCGACATTTCCCGAAAAAGTGCACCTGACGTC 101
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 115 GGGTTCCGGCGACATTTCCCGAAAAAGTGCACCTGACGTC 75
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 3
 AAN10031/C
 ID AAN10031 standard; DNA; 251 BP.
 XX AC AAN10031;
 XX DT 13-AUG-1992 (first entry)
 XX DE Sequence of the pKT241 EcoRI-PstI penicillinase gene fragment.
 XX KW Cloning vehicle; bacterial vector; transformed host; penicillinase;
 XX KW insulin; ds.
 XX OS *Escherichia coli*.
 XX FH Key Location/Qualifiers
 FT misc_feature 1..4
 FT /*tag= a
 FT /label= sticky end
 FT misc_feature 248..251
 FT /*tag= b
 FT /label= sticky end
 XX EP38182-A.
 XX PN 21-OCT-1981.
 XX PD 09-APR-1981; 81EP-00301561.
 XX PF 11-APR-1980; 80US-00139225.
 XX PR (HARD) HARVARD COLLEGE.
 XX PA Gilbert W, Talmadge K;
 XX PI WPI; 1981-80125D/44.
 XX DR P-PSDB; AAP10038.
 XX PS Synthesis of mature protein or polypeptide - by using bacterial host
 PT transformed by cloned vehicle contg. DNA fragment etc.
 XX Example; Fig 2; 34pp; English.
 XX CC The closest identifiable promoter for the penicillinase gene in pKT241
 CC (AAN10031) is located in the region 14 to 20 nucleotides before its
 CC translational start signal. In the examples, the 3' end of pKT241 was
 CC attached to the signal DNA sequence of the DNA fragment (19) for rat
 CC preproinsulin (see AAN10033). The closest identifiable promoter for the
 CC penicillinase gene in pKT218 (AAN10032) is located in the region 14 to 20
 CC nucleotides before its translational start signal. In the examples, the

CC 3' end of pKT218 was attached to the signal DNA sequence of the DNA
 CC fragment (CB6) for rat preproinsulin (see AAN10034)
 XX
 SQ Sequence 251 BP; 74 A; 45 C; 50 G; 82 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 1; Length 251;
 Best Local Similarity 100.0%; Pred. No. 2.3e-21;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGTTATTCTCTCATGCGGATACATATTGAATGTTAGAAAAATAAACAATAG 60
 Db 175 AGGTTATTCTCTCATGCGGATACATATTGAATGTTAGAAAAATAAACAATAG 116

Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 101
 Db 115 GGGTTCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 75

RESULT 4
 AAV31229/c
 ID AAV31229 standard; DNA; 400 BP.
 AC AAV31229;
 XX
 DT 01-OCT-1998 (first entry)
 XX
 KW E. coli J96 pathogenicity island contig #43.
 KW PAI; pathogenicity island; uropathogenic E. coli detection; PAI IV; pher;
 KW PAI V; pher; vaccine; protective immune response; ds.
 OS Escherichia coli.
 XX
 PN WO9822575-A2.
 XX
 PD 28-MAY-1998.
 XX
 PF 21-NOV-1997; 97WO-US021347.
 XX
 PR 22-NOV-1996; 96US-0031626P.
 PR 14-OCT-1997; 97US-0061953P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PA (UYWI-) UNIV WISCONSIN.
 XX
 PI Dillon PJ, Choi GH, Welch RA;
 XX
 DR WPI; 1998-312461/27.
 XX
 PT New isolated uropathogenic E. coli nucleotide sequences - used to develop
 PT products for the detection of pathogenic E. coli and to elicit an immune
 PT response to pathogenic E. coli.
 XX
 PS Claim 21; Page 140-141; 250pp; English.
 XX
 CC This sequence represents a E. coli strain J96 contig containing
 CC pathogenicity island (PAI) sequences, and represents a nucleic acid
 CC molecule of the invention. PAIs are large fragments of DNA which comprise
 CC pathogenicity determinants. The sequences of the invention are taken from
 CC PAI IV and PAI V. PAI IV is located at approximately 64 min (near pher)
 CC on the E. coli chromosome and is greater than 170 kb. PAI V is located at
 CC approximately 94 min (at pher) on the E. coli chromosome and is
 CC approximately 160 kb in size. Antibodies specific to the proteins encoded
 CC by the PAI open reading frames of the invention can be used in kits to
 CC detect uropathogenic E. coli. The proteins are used in vaccines to elicit
 CC a protective immune response in an animal to the uropathogenic E. coli
 CC strain J96
 XX
 SQ Sequence 400 BP; 106 A; 77 C; 91 G; 126 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 2; Length 400;
 Best Local Similarity 100.0%; Pred. No. 2.5e-21;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGTTATTCTCTCATGCGGATACATATTGAATGTTAGAAAAATAAACAATAG 60
 Db 165 AGGTTATTCTCTCATGCGGATACATATTGAATGTTAGAAAAATAAACAATAG 106

Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 101
 Db 105 GGGTTCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 65

RESULT 5
 AAN60624/c
 ID AAN60624 standard; DNA; 456 BP.
 XX
 AC AAN60624;
 XX
 DT 25-MAR-2003 (revised)
 DT 29-OCT-1991 (first entry)
 XX
 KW Plasmid pUG201 sequence encoding beta-urogastrone.
 KW Beta-lactamase signal peptide; pGH54; pGH55; ss.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FH 125..170
 FT promoter /*tag= a
 FT RBS 200..203
 FT CDS /*tag= b
 FT 209..439
 FT /*tag= c
 FT 209..277
 FT /*tag= d
 FT /label= Beta-lactamase signal peptide
 FT 278..436
 FT /*tag= e
 FT /label= Beta-urogastrone
 XX
 PN WO8603779-A.
 XX
 PD 03-JUL-1986.
 XX
 PP 19-DEC-1985; 85WO-JP000696.
 XX
 PR 21-DEC-1984; 84JP-00271206.
 XX
 PA (EART) EARTH CHEM CO LTD.
 PA (OHGA/) OHGAI H.
 XX
 PI Ohgai H, Momota H, Kuwakura T, Kajifusa N, Kitazawa T, Oshiden K;
 XX
 DR WPI; 1986-182911/28.
 DR P-PSDB; AAP60678.
 XX
 PT Recombinant vector for polypeptide secretion - contains signal peptide
 PT sequence directly bonded to peptide-coding sequence.
 XX
 PS Disclosure; Table 4; 79pp; Japanese.
 XX
 CC The plasmid produces secreted beta-urogastrone in a transformed
 CC expression system. Similar plasmids may be constructed where the
 CC secretion signal may be coupled with eg. somatostatin, insulin, growth
 CC hormone, interferon, IL-2, gastric inhibitory peptide, influenza B SA,
 CC epidermal growth factor and thymosine-beta4. (Updated on 25-MAR-2003 to
 CC correct PA field.)
 XX
 SQ Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 1; Length 456;
 Best Local Similarity 100.0%; Pred. No. 2.6e-21;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAATAG 60
 |||
 Db 173 AGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAATAG 114
 |||
 QY 61 GGGTTCCGGCAGACATTTCCCGGAAAGTGCCACCTGACGTC 101
 |||
 Db 113 GGGTTCCGGCAGACATTTCCCGGAAAGTGCCACCTGACGTC 73
 |||

RESULT 6

AAAN71080/c
 ID AAAN71080 standard; DNA; 456 BP.
 AC AAAN71080;
 XX
 DT 25-MAR-2003 (revised)
 DT 10-MAR-2003 (revised)
 DT 13-MAY-1991 (first entry)
 XX
 DE Sequence encoding beta-urogastrone.
 XX
 KW pUGT 150s; beta-UG; ds.
 XX
 OS Escherichia coli.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT promoter 125..170
 FT /*tag= a
 CDS 209..439
 FT /*tag= b
 FT /transl_except= (pos:434..436,aa:Arg)
 FT
 PN JPC2190083-A.
 XX
 XX 20-AUG-1987.
 XX
 XX 14-FEB-1986; 86JJP-00031415.
 XX
 XX 14-FEB-1986; 86JJP-00031415.
 PR
 XX (EART) EARTH SEIYAKU KK.
 PA
 XX WPI; 1987-273761/39.
 DR
 XX Expression vector contg. multiple information units is used to transform host all for increased prodn. of polypeptides.
 PT
 XX Disclosure; Page 553; 34pp; Japanese.
 PS
 XX Sequence encodes beta-urogastrone under the control of a tac promoter. The peptide may be expressed from plasmid pUGT 150s in a transformed E.coli host. The plasmid may carry several separately expressing sequences comprising a tac promoter, SD site, signal peptide, and coding sequence, to produce beta-UG in high yield. (Updated on 10-MAR-2003 to add missing OS field.) (Updated on 25-MAR-2003 to correct PA field.)
 CC
 SQ Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;
 Query Match 100.0%; Score 101; DB 1; Length 456;
 Best Local Similarity 100.0%; Pred. No. 2.6e-21;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAATAG 60
 |||
 Db 173 AGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAATAG 114
 |||
 QY 61 GGGTTCCGGCAGACATTTCCCGGAAAGTGCCACCTGACGTC 101
 |||
 Db 113 GGGTTCCGGCAGACATTTCCCGGAAAGTGCCACCTGACGTC 73
 |||

RESULT 7

AAAN70833/c
 ID AAAN70833 standard; DNA; 456 BP.
 XX
 AC AAAN70833;
 XX
 DT 25-MAR-2003 (revised)
 DT 10-MAR-2003 (revised)
 DT 18-JAN-1991 (first entry)
 XX
 DE Beta-urogastrone sequence.
 XX
 KW Tumour; inosine; DNA probe; ds.
 XX
 OS Unidentified.
 XX
 FH Key Location/Qualifiers
 FT promoter 125..170
 FT /*tag= b
 RBS 200..204
 FT /*tag= c
 FT CDS 209..439
 FT /*tag= a
 FT sig_peptide 209..277
 FT /*tag= d
 XX
 PN JPC2244398-A.
 XX
 XX 24-OCT-1987.
 PD
 XX 16-APR-1986; 86JJP-00087368.
 PF
 XX 16-APR-1986; 86JJP-00087368.
 PR
 XX (SEKI) SEKISUI CHEM IND CO LTD.
 PA
 XX WPI; 1987-339045/48.
 DR
 XX P-PSDB; AAP70505.
 XX
 PT Detection of DNA and/or RNA - by converting to single strand form and using probe contg. labelled inosine deriv.
 PT
 XX Disclosure; Page 11; 11pp; Japanese.
 PS
 XX An example of a sequence detected by a probe consisting of polyinosine, polydeoxyinosine, oligoinosine and/or oligodeoxyinosine is labeled. The ssDNA and probe are hybridized and the existence of DNA in the product is detected. It can be used to detect the presence of malignant tumour.
 CC
 CC (Updated on 10-MAR-2003 to add missing OS field.) (Updated on 25-MAR-2003 to correct PA field.)
 CC
 SQ Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;
 Query Match 100.0%; Score 101; DB 1; Length 456;
 Best Local Similarity 100.0%; Pred. No. 2.6e-21;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAATAG 60
 |||
 Db 173 AGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTAGAAAAATAAACAATAG 114
 |||
 QY 61 GGGTTCCGGCAGACATTTCCCGGAAAGTGCCACCTGACGTC 101
 |||
 Db 113 GGGTTCCGGCAGACATTTCCCGGAAAGTGCCACCTGACGTC 73
 |||
 RESULT 8
 AAAN81765/c
 ID AAAN81765 standard; DNA; 456 BP.
 XX
 AC AAAN81765;
 XX
 DT 25-MAR-2003 (revised)
 DT 13-DEC-1990 (first entry)

```
XX Sequence of HB101 (pGH201) encoding new beta-urogastrone deriv. Met (21),
DE Arg (53).
XX
XX Gastric acid secretion; cell proliferation; hormone; ds.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX CDS 209..277 /*tag= a
XX CDS 278..439 /*tag= b
XX FT /*product= "New beta-urogastrone deriv."
XX
XX JP63012298-A.
XX
XX 19-JAN-1988.
XX
XX 30-JUN-1986; 86JP-00153783.
XX
XX 30-JUN-1986; 86JP-00153783.
XX
XX (EART ) EARTH SEIYAKU KK.
XX
XX WPI; 1988-054638/08.
XX
XX P-PSDB; AAP81349.
XX
XX New beta-urogastrone deriv. - has gastric acid secretion inhibition and
XX proliferation promotion activity.
XX
XX Disclosure; Page 685; 76pp; Japanese.
XX
XX The deriv. has various biological activities such as gastric acid
XX secretion inhibiting action, or cell proliferation promoting action. The
XX deriv. has the same biological or pharmacological activities as beta-
XX urogastrone. It is not susceptible to denaturation by oxidn. and is
XX chemically stable. Deriv. has resistance to proteolytic enzymes such as
XX pepsinase. (Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 101; DB 1; Length 456;
XX Best Local Similarity 100.0%; Pred. No. 2.6e-21;
XX Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAAATAG 60
XX 173 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAAATAG 114
XX
XX 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
XX 113 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 73
XX
XX RESULT 9
XX ABA90413/c
XX ID ABA90413 standard; DNA; 466 BP.
XX
XX AC ABA90413;
XX
XX 12-FEB-2002 (first entry)
XX
XX Drosophila cell cycle progression protein coding sequence #48.
XX
XX Antiproliferative; cytostatic; cardiant; immunosuppressive; meiosis;
XX antiinflammatory; antipsoriatic; dermatological; antifungal; mitosis;
XX antiparasitic; antimalarial; antirheumatic; antiarthritic; cell division;
XX cell cycle progression protein; tumour; proliferative disorder;
XX cardiovascular; autoimmune; dermatological disorder; ds.
XX
XX Drosophila sp.
XX
```

```
PN WO200172774-A2.
XX
XX PD 04-OCT-2001.
XX
XX PF 23-MAR-2001; 2001WO-GB001297.
XX
XX PR 24-MAR-2000; 2000GB-00007268.
XX
XX PA (CYCL-) CYCLACEL LTD.
XX
XX PI Deak P, Glover DM, Midgley C;
XX
XX DR WPI; 2002-055132/07.
XX
XX PT Polynucleotides encoding cell cycle progression proteins, useful for
XX treating a tumor or a proliferative disorder.
XX
XX PS Claim 1; Page 99; 213pp; English.
XX
XX CC The present invention relates to Drosophila cell cycle progression
XX proteins (AAM47572-AAM47608) and their coding sequences (ABA90366-
XX ABA90520). The coding sequences and proteins are useful for identifying a
XX substance capable of affecting the function of the corresponding gene, a
XX inhibiting mitosis and/or meiosis. The cell can also be used in a method for
XX treating a tumour or proliferative disorder, cardiovascular disorders
XX (such as restenosis and cardiomyopathy), autoimmune disorders such as
XX (glomerulonephritis and rheumatoid arthritis), dermatological disorders
XX (such as psoriasis), antiinflammatory, antifungal and antiparasitic
XX disorders (such as malaria)
XX
XX SQ Sequence 466 BP; 135 A; 87 C; 93 G; 150 T; 0 U; 1 Other;
XX
XX Query Match 100.0%; Score 101; DB 6; Length 466;
XX Best Local Similarity 100.0%; Pred. No. 2.6e-21;
XX Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAAATAG 60
XX 280 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAAATAG 221
XX
XX 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
XX 220 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 180
XX
XX RESULT 10
XX AAX21173/c
XX ID AAX21173 standard; DNA; 487 BP.
XX
XX AC AAX21173;
XX
XX 05-MAY-1999 (first entry)
XX
XX Polynucleotide sequence from the genome of Treponema pallidum.
XX
XX Treponema pallidum infection; syphilis; Borrelia infection; animal;
XX enzyme production; ds.
XX
XX Treponema pallidum.
XX
XX OS OS
XX PN PN
XX PD 30-DEC-1998.
XX
XX PF 23-JUN-1998; 98WO-US013041.
XX
XX PR 24-JUN-1997; 97US-0050667P.
XX
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX PI Fraser CM;
XX
```

DR WPI; 1999-081273/07.
XX New isolated Treponema pallidum nucleic acids - used to develop products
PT for the detection, diagnosis, characterisation, prevention and therapy of
PT T. pallidum infections, particularly syphilis.
XX
PS Claim 1; Page 1106; 1150pp; English.
XX
CC AAX20500-21243 represent polynucleotide sequences from the genome of
CC Treponema pallidum. The sequences can be used for detection, diagnosis,
CC characterisation, prevention and therapy for T. pallidum infections,
CC particularly syphilis. They can also be used for detecting diseases
CC related to Borrelia infections in animals, and for the production of
CC biosynthetic products such as enzymes
XX
SQ Sequence 487 BP; 125 A; 127 C; 113 G; 121 T; 0 U; 1 Other;
Query Match 100.0%; Score 101; DB 2; Length 487;
Best Local Similarity 100.0%; Pred. No. 2.6e-21; Indels 0; Gaps 0;
Matches 101; Conservative 0; Mismatches 0;
QY 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTGAAAAATAAACAATAG 60
DB 323 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTGAAAAATAAACAATAG 264
QY 61 GGGTTCGGCGACATTTCCCGAAAAAGTCCACCTGACGTC 101
DB 263 GGGTTCGGCGACATTTCCCGAAAAAGTCCACCTGACGTC 223
RESULT 11
AAX21145/C
ID AAX21149 standard; DNA; 535 BP.
XX
AC AAX21149;
XX
DT 05-MAY-1999 (first entry)
XX
XX Polynucleotide sequence from the genome of Treponema pallidum.
DE Treponema pallidum infection; syphilis; Borrelia infection; animal;
KW enzyme production; ds.
XX Treponema pallidum.
OS
XX WO9859034-A2.
XX
XX 30-DEC-1998.
XX
XX 23-JUN-1998; 98WO-US013041.
XX
XX 24-JUN-1997; 97US-0050667P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Fraser CM;
XX
XX WPI; 1999-081273/07.
XX
XX New isolated Treponema pallidum nucleic acids - used to develop products
PT for the detection, diagnosis, characterisation, prevention and therapy of
PT T. pallidum infections, particularly syphilis.
XX
PS Claim 1; Page 1093; 1150pp; English.
XX
CC AAX20500-21243 represent polynucleotide sequences from the genome of
CC Treponema pallidum. The sequences can be used for detection, diagnosis,
CC characterisation, prevention and therapy for T. pallidum infections,
CC particularly syphilis. They can also be used for detecting diseases
CC related to Borrelia infections in animals, and for the production of
CC biosynthetic products such as enzymes
XX
SQ Sequence 535 BP; 145 A; 108 C; 122 G; 155 T; 0 U; 5 Other;
Query Match 100.0%; Score 101; DB 6; Length 573;
Best Local Similarity 100.0%; Pred. No. 2.7e-21; Indels 0; Gaps 0;
Matches 101; Conservative 0; Mismatches 0;
QY 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTGAAAAATAAACAATAG 60
DB 355 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTGAAAAATAAACAATAG 296

Query Match 100.0%; Score 101; DB 2; Length 535;
Best Local Similarity 100.0%; Pred. No. 2.7e-21; Indels 0; Gaps 0;
Matches 101; Conservative 0; Mismatches 0;
QY 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTGAAAAATAAACAATAG 60
DB 158 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTGAAAAATAAACAATAG 99
QY 61 GGGTTCGGCGACATTTCCCGAAAAAGTCCACCTGACGTC 101
DB 98 GGGTTCGGCGACATTTCCCGAAAAAGTCCACCTGACGTC 58
RESULT 12
ABA90456/c
ID ABA90456 standard; DNA; 573 BP.
XX
AC ABA90456;
XX
DT 12-FEB-2002 (first entry)
XX
XX Drosophila cell cycle progression protein coding sequence #91.
DE
XX
XX Antiproliferative; cytotostatic; cardiant; immunosuppressive; meiosis;
KW antiinflammatory; antipsoriatic; dermatological; antifungal; mitosis;
KW antiparasitic; antimalarial; antirheumatic; antiarthritic; cell division;
KW cell cycle progression protein; tumour; proliferative disorder;
KW cardiovascular; autoimmune; dermatological disorder; ds.
XX
XX Drosophila sp.
XX
XX WO200172774-A2.
XX
XX 04-OCT-2001.
XX
XX 23-MAR-2001; 2001WO-GB001297.
XX
XX 24-MAR-2000; 2000GB-00007268.
XX
XX (CYCL-) CYCLACEL LTD.
XX
XX Deak P, Glover DM, Midgley C;
XX
XX WPI; 2002-055132/07.
XX
XX Polynucleotides encoding cell cycle progression proteins, useful for
PT treating a tumor or a proliferative disorder.
XX
PS Claim 1; Page 144; 213pp; English.
XX
XX The present invention relates to Drosophila cell cycle progression
CC proteins (AAM47572-AAM47608) and their coding sequences (ABA90366-
CC ABA90520). The coding sequences and proteins are useful for identifying a
CC substance capable of affecting the function of the corresponding gene, a
CC substance capable of inhibiting the cell division cycle, or capable of
CC inhibiting mitosis and/or meiosis. They can also be used in a method for
CC treating a tumor or proliferative disorder, cardiovascular disorders
CC (such as restenosis and cardiomyopathy), autoimmune disorders such as
CC (glomerulonephritis and rheumatoid arthritis), dermatological disorders
CC (such as psoriasis), antiinflammatory, antifungal and antiparasitic
CC disorders (such as malaria)
XX
XX
SQ Sequence 573 BP; 154 A; 118 C; 116 G; 184 T; 0 U; 1 Other;
Query Match 100.0%; Score 101; DB 6; Length 573;
Best Local Similarity 100.0%; Pred. No. 2.7e-21; Indels 0; Gaps 0;
Matches 101; Conservative 0; Mismatches 0;
QY 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTGAAAAATAAACAATAG 60
DB 355 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTGAAAAATAAACAATAG 296

QY 61 GGGTTCGGCGACATTTCCCGGAAAGTGCACCTGACGTC 101
|
Db 295 GGGTTCGGCGACATTTCCCGGAAAGTGCACCTGACGTC 255

RESULT 13

ADH58311
ID ADH58311 standard; DNA; 605 BP.

XX AC
XX ADH58311;

XX DT -25-MAR-2004 (first entry)

XX DE Electropherogram of a DNA sequencing reaction using E154 & T422.

XX ds; primer library; extendable oligos; EO; ligation chain reaction; LCR;
KW rolling circle amplification; strand displacement amplification;
KW isothermal DNA amplification; biotechnology; agriculture;
KW medical research; pUC19 plasmid.

XX OS Synthetic.
OS Escherichia coli.

XX PN WO2003093500-A1.

XX PD 13-NOV-2003.

XX PF 24-DEC-2002; 2002WO-AU001763.

XX PR 01-MAY-2002; 2002AU-00002045.

XX PA (NUCL-) NUCLEICS PTY LTD.

XX PI Tillett D, Thomas T;

XX DR WPI; 2004-053046/05.

XX Increasing the affinity of an extendable oligonucleotide (EO) for a
PT target nucleic acid, for providing primers having improved specificity,
PT comprises hybridization of the EO to a template oligonucleotide (TO) and
PT extension of the EO.

XX Example 10; Fig 23; 85pp; English.

XX This invention relates to a novel method for the optimisation of primer
CC libraries. Specifically, it refers to increasing the affinity of short
CC oligonucleotide primers, also known as extendable oligos (EOs), for their
CC template sequences. The present invention describes improved methods for
CC sequencing and the linear and exponential amplification of DNA that can
CC be useful for PCR, RT-PCR, ligation chain reaction (LCR), rolling circle
CC amplification, strand displacement amplification and isothermal DNA
CC amplification. Accordingly, these extendable oligos with improved
CC specificity and affinity are particularly important in fields ranging
CC from biotechnology and agriculture to medical research. This
CC polynucleotide sequence is the electropherogram of a DNA sequencing
CC reaction that used the pUC19 plasmid and E154/T422 oligos, used in an
XX exemplification of the invention.

XX Sequence 605 BP; 159 A; 133 C; 147 G; 148 T; 0 U; 18 Other;

Query Match 100.0%; Score 101; DB 12; Length 605;
Best Local Similarity 100.0%; Pred. No. 2.8e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATGTCTCATGACGGGATACATATTGTAATGATTATTTAGAAAAATAAACAATAG 60
|
Db 259 AGGGTTATGTCTCATGACGGGATACATATTGTAATGATTATTTAGAAAAATAAACAATAG 318

QY 61 GGGTTCGGCGACATTTCCCGGAAAGTGCACCTGACGTC 101
|
Db 319 GGGTTCGGCGACATTTCCCGGAAAGTGCACCTGACGTC 359

RESULT 14

AAS30560/C

ID AAS30560 standard; DNA; 776 BP.

XX AC AAS30560;

XX DT 21-NOV-2001 (first entry)

XX DE DNA encoding novel prostate gland antigen, Seq ID No 418.

XX Human; neurotropic; neuroprotective; cytostatic; antiparkinsonian;
KW antianaeic; dermatological; immunosuppressive; antiinflammatory;
KW antiarthritic; antirheumatic; virucide; hepatotropic; nephrotropic;
KW osteopathic; prostate gland; prostatitis; adenocarcinoma; hair loss;
KW prostatic; malacoplakia; adenocarcinoma; benign prostatic hypertrophy;
KW hyperplasia; carcinoma; prostate neoplastic disorder; skin aging;
KW reproductive system disorder; autoimmune disorder; urinary system;
KW systemic lupus erythematosus; rheumatoid arthritis; cardiovascular;
KW blood-related disorder; hyperproliferative disorder; respiratory;
KW neurological disorder; endocrine disorder; inflammatory disorder;
KW liver disorder; wound healing; food preservative; ds.

XX OS Homo sapiens.

XX PN WO200155447-A1.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001330.

XX PR 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 18-AUG-2000; 2000US-0225759P.

PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226688P.

PR 22-AUG-2000; 2000US-0227182P.

PR 23-AUG-2000; 2000US-0227009P.

PR 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.

PR 01-SEP-2000; 2000US-0229343P.

PR 01-SEP-2000; 2000US-0229344P.

PR 01-SEP-2000; 2000US-0229345P.

PR 05-SEP-2000; 2000US-0229509P.

PR 05-SEP-2000; 2000US-0229513P.

PR 06-SEP-2000; 2000US-0230437P.

PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 08-SEP-2000; 2000US-0232082P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241856P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.

PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-476223/51.

XX Novel isolated prostate gland related polypeptide useful for diagnosis
PT and treatment of disorders of prostate such as prostatodystonia,
PT prostatosis, prostatitis, benign prostatic hypertrophy and malacoplakia.

PS Claim 1; SEQ ID NO 418; 512pp; English.

XX The invention relates to novel isolated prostate gland related nucleic
CC acids (I) and polypeptides (II). (I) and (II) are useful for diagnosis,
CC prognosis, prevention, and/or treatment of diseases and/or disorders of
CC the prostate such as acute non-bacterial prostatitis, chronic non-
CC bacterial prostatitis, acute bacterial prostatitis, prostatodystonia,
CC prostatosis, granulomatous prostatitis, malacoplakia, benign prostatic
CC hypertrophy or hyperplasia, and prostate neoplastic disorders, including
CC adenocarcinomas, transitional cell carcinomas, ductal carcinomas, and
CC squamous cell carcinomas. (I), (II) and antibody to (II) are useful for
CC diagnosing and treating reproductive system disorders (Paget's disease),
CC autoimmune disorders (systemic lupus erythematosus, rheumatoid
CC arthritis), blood-related disorders (sickle cell anaemia),
CC hyperproliferative disorders, urinary system disorders
CC (glomerulonephritis), cardiovascular disorders (arrhythmias), respiratory
CC disorders, musculoskeletal system disorders, neural activity and
CC neurological disorders (Alzheimer's disease and Parkinson's disease),
CC endocrine disorders (Addison's disease), gastrointestinal disorders
CC (inflammatory disorders), liver disorders (biliary liver cirrhosis),
CC pancreatic and gall bladder disorders, disorders of the large intestine,
CC developmental and inherited disorders, diseases at the cellular level,
CC and wound healing and epithelial cell proliferation. (I) or (II) is
CC useful to prevent skin aging, for preventing hair loss, to maintain
CC organs before transplantation, and as food additive or preservative.

Query Match 100.0%; Score 101; DB 4; Length 776;

Best Local Similarity 100.0%; Pred. No. 2.9e-21;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGAATGTTATTTAGAAAAATAAATAG 60

Db 546 AGGGTTATTGTCATGAGCGGATACATATTGAATGTTATTTAGAAAAATAAATAG 487

Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCCACTGACGTC 101

Db 486 GGGTTCGCGCACATTTCCCGAAAGTGCCACTGACGTC 446

RESULT 15

AAS27819/c

ID AAS27819 standard; DNA; 776 BP.

XX AAS27819;
XX
XX
DT
XX
XX
DE
XX
KW DNA encoding novel signal transduction pathway protein, Seq ID 1479.
KW Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
KW anti-inflammatory; anti-HIV; antibacterial; anti-inflammatory; cancer;
KW immune system disorder; rheumatoid arthritis; inflammatory condition;
KW organ transplant rejection; infection; hepatitis C; blood disorder;
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
KW reproductive system; gastrointestinal; liver disorder; AIDS; ds;
KW acquired immune deficiency syndrome.
OS
XX Homo sapiens.
XX
XX WO200154733-A1.
XX
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001312.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205151P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0228287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 12-SEP-2000; 2000US-0232081P.
PR 14-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 26-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.

PR	17-NOV-2000;	2000US-0249297P.
PR	17-NOV-2000;	2000US-0249299P.
PR	17-NOV-2000;	2000US-0249300P.
PR	01-DEC-2000;	2000US-0250160P.
PR	01-DEC-2000;	2000US-0250391P.
PR	05-DEC-2000;	2000US-0251030P.
PR	05-DEC-2000;	2000US-0251988P.
PR	05-DEC-2000;	2000US-0256719P.
PR	06-DEC-2000;	2000US-0251479P.
PR	08-DEC-2000;	2000US-0251856P.
PR	08-DEC-2000;	2000US-0251868P.
PR	08-DEC-2000;	2000US-0251869P.
PR	08-DEC-2000;	2000US-0251989P.
PR	08-DEC-2000;	2000US-0251990P.
PR	11-DEC-2000;	2000US-0254097P.
XX	05-JAN-2001;	2001US-0259678P.
XX		
PA	(HUMA-) HUMAN GENOME SCI INC.	
XX		
PI	Rosen CA, Barash SC, Ruben SM;	
XX		
DR	WPI; 2001-465460/50.	
XX		
XX		
PT	Novel polypeptides useful for diagnosing, treating, preventing and/or	
PT	prognosing disorders related to the proteins, including cancers, immune	
PT	disorders and neuronal disorders.	
XX		
PS	Claim 1; SEQ ID NO 1479; 880pp; English.	
XX		
CC	The invention relates to novel isolated polypeptides (I), and	
CC	polynucleotides (II). (I) and the antibody to (I) are useful for	
CC	diagnosing, preventing and treating diseases including immune system	
CC	disorders (e.g. congenital and acquired immunodeficiencies, autoimmune	
CC	disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ	
CC	transplant rejections and graft versus host disease, infectious diseases	
CC	(e.g. hepatitis C) bleeding disorders, haemoglobin abnormalities and	
CC	other blood-related disorders (sickle cell anaemia), myeloproliferative	
CC	disorders, primary haematopoietic disorders, hyperproliferative disorders	
CC	(e.g. Gaucher's disease and cancer), neurodegenerative disorders (e.g.	
CC	Alzheimer's disease, Parkinson's disease), chromosomal abnormalities	
CC	(Down syndrome), ischaemic injury (e.g. stroke), renal disorders (e.g.	
CC	glomerulonephritis), cardiovascular disorders (e.g. arrhythmia),	
CC	respiratory disorders, dermatological disorders, in wound healing,	
CC	epithelial cell proliferation, endocrine disorders (e.g. Addison's	
CC	disease), reproductive system disorders, gastrointestinal disorder	
CC	(inflammatory disorders), liver disorders (cirrhosis), as stimulators of	
CC	B-cell responsiveness to pathogens, activators of T-cells, to induce	
CC	higher affinity antibodies, and as a means to induce tumour proliferation	
CC	in pathologies e.g. acquired immune deficiency syndrome (AIDS). AAS26976-	
CC	AAS27850 represent novel signal transduction pathway protein coding	
CC	sequences and PCR primers of the invention	
XX		
SQ	Sequence 776 BP; 184 A; 209 C; 181 G; 200 T; 0 U; 2 Other;	
	Query Match 100.0%; Score 101; DB 4; Length 776;	
	Best Local Similarity 100.0%; Pred. No. 2.9e-21;	
	Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
QY	1 AGGGTTATTGTCATGAGCGGTACATATTTGAATGATTATTAGAAAAATAACAATAAG 60	
Db	546 AGGGTTATTGTCATGAGCGGTACATATTTGAATGATTATTAGAAAAATAACAATAAG 487	
QY	61 GGGTTCGCGGCACATTTCCCCGAAAAGTGCCACCTGACGTC 101	
Db	486 GGGTTCGCGGCACATTTCCCCGAAAAGTGCCACCTGACGTC 446	

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 05:15:57 ; Search time 961.667 Seconds
(without alignments)

3997.736 Million cell updates/sec

Title: US-09-482-682-64_COPY_7131_7231

Perfect score: 101

Sequence: 1 aggggtattgtctcatgagc.....gaaagtgcacatgacgtc 101

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hic:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gss1:*

9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	101	100.0	300	4	BM078095
C 2	101	100.0	300	5	BU963956
C 3	101	100.0	300	5	BU964094
C 4	101	100.0	309	9	FR0009140
C 5	101	100.0	391	1	AL597149
C 6	101	100.0	414	9	CC819240
C 7	101	100.0	417	4	BJ684174
C 8	101	100.0	491	9	CC819923
C 9	101	100.0	495	4	BI805285
C 10	101	100.0	495	9	CC818374
C 11	101	100.0	496	9	CC818523
C 12	101	100.0	503	9	CC819854
C 13	101	100.0	515	9	CC817752
C 14	101	100.0	518	9	CC817128
C 15	101	100.0	519	9	CC817162
C 16	101	100.0	519	9	CC817796
C 17	101	100.0	521	9	CC819067
C 18	101	100.0	533	9	CC819841
C 19	101	100.0	542	9	CC816892
C 20	101	100.0	550	7	CF766622
C 21	101	100.0	551	9	CC816905
C 22	101	100.0	554	9	CC819058
C 23	101	100.0	563	9	CC819270
C 24	101	100.0	566	9	CC816848

C 25	101	100.0	566	9	CC820024
C 26	101	100.0	567	9	CC817070
C 27	101	100.0	568	9	CC818640
C 28	101	100.0	571	9	CC816954
C 29	101	100.0	571	9	CC818423
C 30	101	100.0	580	9	CC819098
C 31	101	100.0	582	1	AL694813
C 32	101	100.0	583	9	CC817633
C 33	101	100.0	583	9	CC818436
C 34	101	100.0	586	9	CC816883
C 35	101	100.0	588	9	CC817788
C 36	101	100.0	588	9	CC818340
C 37	101	100.0	589	9	CC817595
C 38	101	100.0	590	9	CC819754
C 39	101	100.0	592	9	CC817679
C 40	101	100.0	592	9	CC818508
C 41	101	100.0	593	9	CC816942
C 42	101	100.0	593	9	CC817699
C 43	101	100.0	594	9	CC818287
C 44	101	100.0	594	9	CC818422
C 45	101	100.0	595	9	CC816929

ALIGNMENTS

RESULT 1

BM078095/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BM078095 300 bp mRNA linear EST 30-NOV-2001
83374 Hebeloma cylindrosporum functional cDNA library Hebeloma
cylindrosporum cDNA 5', mRNA sequence.
BM078095
BM078095.1 GI:17157967
EST
Hebeloma cylindrosporum
Hebeloma cylindrosporum
Rukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
Agaricales; Cortinariaceae; Hebeloma.
Wipf D., Benjidia M., Tegeder M. and Frommer W.B.
Construction of a functional cDNA library from the ectomycorrhizal
fungus Hebeloma cylindrosporum
Unpublished (2001)
Contact: Wipf D
ZMBP - Center for Molecular Biology of Plants
University of Tuebingen
Auf der Morgenstelle 1, 72076 Tuebingen, Germany
Tel: 49 7071 2976160
Fax: 49 7071 293287
Email: daniel.wipf@zmbp.uni-tuebingen.de
PCR Primers
FORWARD: pDR196 5' primer (PMA 5')
High quality sequence stop: 300
POLYA=No.

FEATURES

source

1. .300

/organism="Hebeloma cylindrosporum"

/mol_type="mRNA"

/strain="H1"

/db_xref="taxon:76867"

/tissue_type="Mycelia"

/lab_host="E. coli XLI-Blue"

/clone_lib="Hebeloma cylindrosporum functional cDNA library"

/notes="vector: pDR 196 (unpublished); Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Query Match

Best Local Similarity

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

100.0%; Score 101; DB 4; Length 300;

100.0%; Pred. No. 8.1e-19;

AGGGTATTGCTCATGCGGATACATATTGTAATGCTATTAGAAAAATAACAATAG 60

```

|||||
174 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAAATAG 115
|||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGACGTC 101
|||||
Db 114 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGACGTC 74
|||||

RESULT 2
BU963956/c
LOCUS EST88 Hebeloma cylindrosporum functional cDNA library Hebeloma
DEFINITION cylindrosporum cDNA, mRNA sequence.
ACCESSION BU963956
VERSION BU963956.1 GI:24204753
KEYWORDS EST.
SOURCE Hebeloma cylindrosporum
ORGANISM Hebeloma cylindrosporum
Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
Agaricales; Cortinariaceae; Hebeloma.
REFERENCE 1 (bases 1 to 300)
AUTHORS Wipf,D., Benjdia,M., Rikirsch,E., Zimmermann,S., Tegeder,M. and
Frommer,W.B.
TITLE A Hebeloma cylindrosporum expression cDNA library for suppression
cloning in yeast mutants, complementation of a yeast his4 mutant
and EST analysis
JOURNAL Unpublished (2002)
COMMENT Contact: Wipf D
ZMBP - Center for Molecular Biology of Plants
University of Tuebingen
Auf der Morgenstelle 1, 72076 Tuebingen, Germany
Tel: 49 7071 2976160
Fax: 49 7071 293287
Email: daniel.wipf@zmbp.uni-tuebingen.de
homolog to Enterobacteria phage f1 ; ampicillinase (1e-10) .
FEATURES
source
1..300
/organism="Hebeloma cylindrosporum"
/mol_type="mRNA"
/strain="H1"
/db_xref="taxon:76867"
/tissue_type="Mycelia"
/lab_host="E. coli XL1-Blue"
/clone_lib="Hebeloma cylindrosporum functional cDNA
library"
/notes="Vector: pDR 196 (unpublished); Site_1: EcoRI;
Site_2: XhoI"

ORIGIN
Query Match 100.0%; Score 101; DB 5; Length 300;
Best Local Similarity 100.0%; Pred. No. 8.1e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAAATAG 60
|||||
Db 170 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAAATAG 111
|||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGACGTC 101
|||||
Db 110 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGACGTC 70
|||||

RESULT 3
BU964094/c
LOCUS EST226 Hebeloma cylindrosporum functional cDNA library Hebeloma
DEFINITION cylindrosporum cDNA, mRNA sequence.
ACCESSION BU964094
VERSION BU964094.1 GI:24204891
KEYWORDS EST.
SOURCE Hebeloma cylindrosporum
ORGANISM Hebeloma cylindrosporum
Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
Agaricales; Cortinariaceae; Hebeloma.
REFERENCE 1 (bases 1 to 300)
AUTHORS Wipf,D., Benjdia,M., Rikirsch,E., Zimmermann,S., Tegeder,M. and
Frommer,W.B.
TITLE A Hebeloma cylindrosporum expression cDNA library for suppression
cloning in yeast mutants, complementation of a yeast his4 mutant
and EST analysis
JOURNAL Unpublished (2002)
COMMENT Contact: Wipf D
ZMBP - Center for Molecular Biology of Plants
University of Tuebingen
Auf der Morgenstelle 1, 72076 Tuebingen, Germany
Tel: 49 7071 2976160
Fax: 49 7071 293287
Email: daniel.wipf@zmbp.uni-tuebingen.de
homolog to Enterobacteria phage f1 ; ampicillinase (1e-10) .
FEATURES
source
1..300
/organism="Hebeloma cylindrosporum"
/mol_type="mRNA"
/strain="H1"
/db_xref="taxon:76867"
/tissue_type="Mycelia"
/lab_host="E. coli XL1-Blue"
/clone_lib="Hebeloma cylindrosporum functional cDNA
library"
/notes="Vector: pDR 196 (unpublished); Site_1: EcoRI;
Site_2: XhoI"

ORIGIN
Query Match 100.0%; Score 101; DB 5; Length 300;
Best Local Similarity 100.0%; Pred. No. 8.1e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAAATAG 60
|||||
Db 170 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAAATAG 111
|||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGACGTC 101
|||||
Db 110 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGACGTC 70
|||||

RESULT 4
FR0009140
LOCUS F.rubripes GSS sequence, clone 010H20aC4, genomic survey sequence.
DEFINITION F.rubripes GSS sequence, clone 010H20aC4, genomic survey sequence.
ACCESSION AL000426
VERSION AL000426.1 GI:2438278
KEYWORDS GSS; genome survey sequence.
SOURCE Takifugu rubripes (Fugu rubripes)
ORGANISM Takifugu rubripes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Takifugu.
REFERENCE 1
AUTHORS Elgar,G., Clark,M.S., Meek,S., Smith,S., Warner,S., Edwards,Y.J.,
Bouchireb,N., Cottage,A., Yeo,G.S., Umrانيا,Y., Williams,G. and
Brenner,S.
TITLE Generation and analysis of 25 Mb of genomic DNA from the pufferfish
Fugu rubripes by sequence scanning
JOURNAL Genome Res. 9 (10), 960-971 (1999)
MEDLINE 99455097
PUBMED 10523524
REFERENCE 2 (bases 1 to 309)
AUTHORS Elgar,G., Clark,M., Smith,S., Meek,S., Warner,S., Umrانيا,Y.,
Williams,G. and Brenner,S.
TITLE Direct Submission
JOURNAL Submitted (09-SEP-1997) MRC Human Genome Mapping Project Resource
Centre Hinxton, Cambridge, CB10 1SB. Email: biohelp@hgm.mrc.ac.uk
COMMENT Vector: pBluescript II KS
V_type: phagemid

```

```

Agaricales; Cortinariaceae; Hebeloma.
REFERENCE 1 (bases 1 to 300)
AUTHORS Wipf,D., Benjdia,M., Rikirsch,E., Zimmermann,S., Tegeder,M. and
Frommer,W.B.
TITLE A Hebeloma cylindrosporum expression cDNA library for suppression
cloning in yeast mutants, complementation of a yeast his4 mutant
and EST analysis
JOURNAL Unpublished (2002)
COMMENT Contact: Wipf D
ZMBP - Center for Molecular Biology of Plants
University of Tuebingen
Auf der Morgenstelle 1, 72076 Tuebingen, Germany
Tel: 49 7071 2976160
Fax: 49 7071 293287
Email: daniel.wipf@zmbp.uni-tuebingen.de
no homolog below 1e-10.
FEATURES
Location/Qualifiers
source
1..300
/organism="Hebeloma cylindrosporum"
/mol_type="mRNA"
/strain="H1"
/db_xref="taxon:76867"
/tissue_type="Mycelia"
/lab_host="E. coli XL1-Blue"
/clone_lib="Hebeloma cylindrosporum functional cDNA
library"
/notes="Vector: pDR 196 (unpublished); Site_1: EcoRI;
Site_2: XhoI"

ORIGIN
Query Match 100.0%; Score 101; DB 5; Length 300;
Best Local Similarity 100.0%; Pred. No. 8.1e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAAATAG 60
|||||
Db 170 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTAGAAAAATAACAAATAG 111
|||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGACGTC 101
|||||
Db 110 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGACGTC 70
|||||

RESULT 4
FR0009140
LOCUS F.rubripes GSS sequence, clone 010H20aC4, genomic survey sequence.
DEFINITION F.rubripes GSS sequence, clone 010H20aC4, genomic survey sequence.
ACCESSION AL000426
VERSION AL000426.1 GI:2438278
KEYWORDS GSS; genome survey sequence.
SOURCE Takifugu rubripes (Fugu rubripes)
ORGANISM Takifugu rubripes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Takifugu.
REFERENCE 1
AUTHORS Elgar,G., Clark,M.S., Meek,S., Smith,S., Warner,S., Edwards,Y.J.,
Bouchireb,N., Cottage,A., Yeo,G.S., Umrانيا,Y., Williams,G. and
Brenner,S.
TITLE Generation and analysis of 25 Mb of genomic DNA from the pufferfish
Fugu rubripes by sequence scanning
JOURNAL Genome Res. 9 (10), 960-971 (1999)
MEDLINE 99455097
PUBMED 10523524
REFERENCE 2 (bases 1 to 309)
AUTHORS Elgar,G., Clark,M., Smith,S., Meek,S., Warner,S., Umrانيا,Y.,
Williams,G. and Brenner,S.
TITLE Direct Submission
JOURNAL Submitted (09-SEP-1997) MRC Human Genome Mapping Project Resource
Centre Hinxton, Cambridge, CB10 1SB. Email: biohelp@hgm.mrc.ac.uk
COMMENT Vector: pBluescript II KS
V_type: phagemid

```

```

PRIMER: KS
DESCR:
One pass dye-terminator sequencing of cosmid cloned genomic
sequence.
FEATURES
    source
        1..309
            /organism="Takifugu rubripes"
            /mol_type="genomic DNA"
            /db_xref="taxon:31033"
            /clone="010H20aC4"
            /clone_lib="cosmid 010H20"
ORIGIN
    Query Match      100.0%; Score 101; DB 9; Length 309;
    Best Local Similarity 100.0%; Pred. No. 8.2e-19;
    Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGTTATTGCTCATGAGCGGATACATATTTGAATGTTATTTAGAAAAATAAACAAATAG 60
    |||||||
Db 39 AGGTTATTGCTCATGAGCGGATACATATTTGAATGTTATTTAGAAAAATAAACAAATAG 98
    |||||||
QY 61 GGGTTCGGCGCACATTTCCCGGAAAGTGCCACCTGACGTC 101
    |||||||
Db 99 GGGTTCGGCGCACATTTCCCGGAAAGTGCCACCTGACGTC 139
    |||||||

RESULT 5
AL597149      391 bp      mRNA      linear      EST 04-SEP-2003
LOCUS
DEFINITION
DKFP313J1611_r1 313 (synonym: hlcc2) Homo sapiens cDNA clone
AL597149
VERSION
DKFP313J1611 5', mRNA sequence.
AL597149.1 GI:15154845
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS
Koehler, K., Beyer, A., Mewes, W., Weil, B. and Wiemann, S.
TITLE
EST (Koehler, K., Beyer, A., Mewes, W., Weil, B. and Wiemann, S.)
JOURNAL
Unpublished (1999)
COMMENT
Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert.
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by BMPZ (Biomedical Research Center at the Charite,
Berlin/Germany) within the cDNA sequencing consortium of the German
Genome Project.
No sl sequence available.
This clone (DKFP313J1611) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
    source
        1..391
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="DKFP313J1611"
            /dev_stage="adult"
            /lab_host="DH10B"
            /clone_lib="313 (synonym: hlcc2)"
            /note="vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB;
            cDNA-collection"
ORIGIN
    Query Match      100.0%; Score 101; DB 1; Length 391;
    Best Local Similarity 100.0%; Pred. No. 8.3e-19;
    Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGTTATTGCTCATGAGCGGATACATATTTGAATGTTATTTAGAAAAATAAACAAATAG 60
    |||||||
Db 1 AGGTTATTGCTCATGAGCGGATACATATTTGAATGTTATTTAGAAAAATAAACAAATAG 60
    |||||||

PRIMER: KS
DESCR:
One pass dye-terminator sequencing of cosmid cloned genomic
sequence.
FEATURES
    source
        1..309
            /organism="Takifugu rubripes"
            /mol_type="genomic DNA"
            /db_xref="taxon:31033"
            /clone="010H20aC4"
            /clone_lib="cosmid 010H20"
ORIGIN
    Query Match      100.0%; Score 101; DB 9; Length 309;
    Best Local Similarity 100.0%; Pred. No. 8.2e-19;
    Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGTTATTGCTCATGAGCGGATACATATTTGAATGTTATTTAGAAAAATAAACAAATAG 60
    |||||||
Db 39 AGGTTATTGCTCATGAGCGGATACATATTTGAATGTTATTTAGAAAAATAAACAAATAG 98
    |||||||
QY 61 GGGTTCGGCGCACATTTCCCGGAAAGTGCCACCTGACGTC 101
    |||||||
Db 99 GGGTTCGGCGCACATTTCCCGGAAAGTGCCACCTGACGTC 139
    |||||||

RESULT 6
CC819240/c      414 bp      DNA      linear      GSS 17-JUL-2003
LOCUS
DEFINITION
100005D19R Oxytricha plasmid UUGC10 library Sterkiella
histriomuscorum genomic clone UUGC100005D19 R, genomic survey
sequence.
CC819240
VERSION
CC819240.1 GI:32899308
KEYWORDS
GSS.
SOURCE
Sterkiella histriomuscorum (Oxytricha trifallax)
ORGANISM
Sterkiella histriomuscorum
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
Stichotrichida; Oxytrichidae; Sterkiella.
1 (bases 1 to 414)
Dunn, D., Doak, T., Herrick, G. and Weiss, R.
Paired end reads from plasmid inserts of Oxytricha trifallax
macronuclear chromosomes
Unpublished (2003)
JOURNAL
COMMENT
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Plate: 0005 row: D column: 19
Seq primer: CACACGAGAAACAGCTAIGACC
Class: plasmid ends
High quality sequence stop: 414.
FEATURES
    source
        1..414
            /organism="Sterkiella histriomuscorum"
            /mol_type="genomic DNA"
            /db_xref="taxon:94289"
            /clone="UUGC100005D19"
            /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
            /clone_lib="Oxytricha plasmid UUGC10 library"
            /note="Vector: FWD42nv; Purified macronuclear chromosomal
            DNA from Oxytricha trifallax was blunt end-repaired with
            T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
            oligonucleotides were ligated to the blunt ends in high
            molar excess. Vector DNA was prepared from a derivative of
            pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
            derivative of plasmid R1. The vector was ligated with
            adaptors complementary to the insert adaptors and
            purified. The sheared, adapted mouse DNA was annealed to
            adapted vector DNA, and transformed into
            chemically-competent E. Coli XL10-Gold (Stratagene) cells
            and selected for ampicillin resistance."
ORIGIN
    Query Match      100.0%; Score 101; DB 9; Length 414;
    Best Local Similarity 100.0%; Pred. No. 8.3e-19;
    Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGTTATTGCTCATGAGCGGATACATATTTGAATGTTATTTAGAAAAATAAACAAATAG 60
    |||||||
Db 414 AGGTTATTGCTCATGAGCGGATACATATTTGAATGTTATTTAGAAAAATAAACAAATAG 355
    |||||||
QY 61 GGGTTCGGCGCACATTTCCCGGAAAGTGCCACCTGACGTC 101
    |||||||
Db 354 GGGTTCGGCGCACATTTCCCGGAAAGTGCCACCTGACGTC 314
    |||||||

```

```

RESULT 7
BJ684174/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
ORIGIN
Query Match
Best Local Similarity
Matches
101; Conservative
0; Mismatches
0; Indels
0; Gaps
0;
Qy
1
AGGGTTATTGTCATGAGCGGATACATATTTGAATGTTATTTAGAAAAATAACAATAG
60
Db
129
AGGGTTATTGTCATGAGCGGATACATATTTGAATGTTATTTAGAAAAATAACAATAG
70
Qy
61
GGGTTCCGGCCACATTTCCCGAAAGTCGCCACCTGACGTC
101
Db
69
GGGTTCCGGCCACATTTCCCGAAAGTCGCCACCTGACGTC
29
RESULT 8
CC819923/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```



```

/tissue_type="Stem"
/dev_stages="3-5 leaf stage"
/clone_lib="Stem library from Oryza sativa (3-5 leaf
stage)"
/note="Vector: pSport2"

ORIGIN
Query Match      100.0%; Score 101; DB 4; Length 495;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAATAAACAATAG 60
    |||||||
Db 62 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAATAAACAATAG 121

Qy 61 GGGTTCGCGCACATTTCCCGGAAAAGTGCACCTGACGTC 101
    |||||||
Db 122 GGGTTCGCGCACATTTCCCGGAAAAGTGCACCTGACGTC 162

RESULT 10
CC818374/c
LOCUS
DEFINITION
CC818374 495 bp DNA linear GSS 17-JUL-2003
100004807R Oxytricha plasmid UUGC100004807 R, genomic survey
sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Sterkiella histriomuscorum (Oxytricha trifallax)
Sterkiella histriomuscorum
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
Stichotrichida; Oxytrichidae; Sterkiella.
1 (bases 1 to 495)
Dunn, D., Doak, T., Herrick, G. and Weiss, R.
Paired end reads from plasmid inserts of Oxytricha trifallax
macronuclear chromosomes
Unpublished (2003)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Plate: 0004 row: B column: 07
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 495.
Location/Qualifiers
1..495
/organism="Sterkiella histriomuscorum"
/mol_type="genomic DNA"
/db_xref="taxon:94289"
/clone="UUGC100004807"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Oxytricha plasmid UUGC10 library"
/note="Vector: PWD42nv; Purified macronuclear chromosomal
DNA from Oxytricha trifallax was blunt end-repaired with
T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
oligonucleotides were ligated to the blunt ends in high
molar excess. Vector DNA was prepared from a derivative of
PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
derivative of plasmid R1. The vector was ligated with
adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
chemically-competent E. Coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

FEATURES
source
1..495
/organism="Sterkiella histriomuscorum"
/mol_type="genomic DNA"
/db_xref="taxon:94289"
/clone="UUGC100004807"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Oxytricha plasmid UUGC10 library"
/note="Vector: PWD42nv; Purified macronuclear chromosomal
DNA from Oxytricha trifallax was blunt end-repaired with
T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
oligonucleotides were ligated to the blunt ends in high
molar excess. Vector DNA was prepared from a derivative of
PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
derivative of plasmid R1. The vector was ligated with
adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
chemically-competent E. Coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

ORIGIN
Query Match      100.0%; Score 101; DB 9; Length 495;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAATAAACAATAG 60
    |||||||
Db 391 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAATAAACAATAG 332

Qy 61 GGGTTCGCGCACATTTCCCGGAAAAGTGCACCTGACGTC 101
    |||||||

```

```

Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAATAAACAATAG 60
    |||||||
Db 392 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAATAAACAATAG 333

Qy 61 GGGTTCGCGCACATTTCCCGGAAAAGTGCACCTGACGTC 101
    |||||||
Db 332 GGGTTCGCGCACATTTCCCGGAAAAGTGCACCTGACGTC 292

RESULT 11
CC818523/c
LOCUS
DEFINITION
CC818523 496 bp DNA linear GSS 17-JUL-2003
10004L13R Oxytricha plasmid UUGC10 library Sterkiella
histriomuscorum genomic clone UUGC100004L13 R, genomic survey
sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Sterkiella histriomuscorum (Oxytricha trifallax)
Sterkiella histriomuscorum
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
Stichotrichida; Oxytrichidae; Sterkiella.
1 (bases 1 to 496)
Dunn, D., Doak, T., Herrick, G. and Weiss, R.
Paired end reads from plasmid inserts of Oxytricha trifallax
macronuclear chromosomes
Unpublished (2003)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Plate: 0004 row: L column: 13
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 496.
Location/Qualifiers
1..496
/organism="Sterkiella histriomuscorum"
/mol_type="genomic DNA"
/db_xref="taxon:94289"
/clone="UUGC100004L13"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Oxytricha plasmid UUGC10 library"
/note="Vector: PWD42nv; Purified macronuclear chromosomal
DNA from Oxytricha trifallax was blunt end-repaired with
T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
oligonucleotides were ligated to the blunt ends in high
molar excess. Vector DNA was prepared from a derivative of
PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
derivative of plasmid R1. The vector was ligated with
adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
chemically-competent E. Coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

ORIGIN
Query Match      100.0%; Score 101; DB 9; Length 496;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAATAAACAATAG 60
    |||||||
Db 391 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAATAAACAATAG 332

Qy 61 GGGTTCGCGCACATTTCCCGGAAAAGTGCACCTGACGTC 101
    |||||||

```


Stichotrichida; Oxytrichidae; Sterkiella.
 1 (bases 1 to 518)
 Dunn, D., Doak, T., Herrick, G. and Weiss, R.
 Paired end reads from plasmid inserts of Oxytricha trifallax
 macronuclear chromosomes
 Unpublished (2003)
 Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Plate: 0002 row: D column: 21
 Seq primer: CACACAGGAAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 518.

FEATURES
 source
 1..518
 /location/Qualifiers
 /organism="Sterkiella histriomuscorum"
 /mol_type="genomic DNA"
 /db_xref="taxon:94289"
 /clone="UUGC100002D21"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /clone_lib="Oxytricha plasmid UUGC10 library"
 /note="Vector: PWD42nv; Purified macronuclear chromosomal
 DNA from Oxytricha trifallax was blunt end-repaired with
 T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
 oligonucleotides were ligated to the blunt ends in high
 molar excess. Vector DNA was prepared from a derivative of
 pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
 derivative of plasmid R1. The vector was ligated with
 adaptors complementary to the insert adaptors and
 purified. The sheared, adapted mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. Coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

Query Match 100.0%; Score 101; DB 9; Length 518;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTTAGAAAAATAACAATAG 60
 |||||
Db 410 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTTAGAAAAATAACAATAG 351
 |||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
 |||||
Db 350 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 310
 |||||

RESULT 15
CC817162/c
LOCUS 519 bp DNA linear GSS 17-JUL-2003
DEFINITION Oxytricha plasmid UUGC10 library Sterkiella
 histriomuscorum genomic clone UUGC100002J19 R, genomic survey
 sequence.
CC817162
VERSION CC817162.1 GI:32896449
KEYWORDS GSS.
SOURCE Sterkiella histriomuscorum (Oxytricha trifallax)
ORGANISM Sterkiella histriomuscorum
 Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichida;
 Stichotrichida; Oxytrichidae; Sterkiella.
REFERENCE 1 (bases 1 to 519)
AUTHORS Dunn, D., Doak, T., Herrick, G. and Weiss, R.
TITLE Paired end reads from plasmid inserts of Oxytricha trifallax
 macronuclear chromosomes
 Unpublished (2003)
JOURNAL Contact: Robert B. Weiss
COMMENT University of Utah Genome Center

University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Plate: 0002 row: J column: 19
 Seq primer: CACACAGGAAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 519.

FEATURES
 source
 1..519
 /location/Qualifiers
 /organism="Sterkiella histriomuscorum"
 /mol_type="genomic DNA"
 /db_xref="taxon:94289"
 /clone="UUGC100002J19"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /clone_lib="Oxytricha plasmid UUGC10 library"
 /note="Vector: PWD42nv; Purified macronuclear chromosomal
 DNA from Oxytricha trifallax was blunt end-repaired with
 T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
 oligonucleotides were ligated to the blunt ends in high
 molar excess. Vector DNA was prepared from a derivative of
 pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
 derivative of plasmid R1. The vector was ligated with
 adaptors complementary to the insert adaptors and
 purified. The sheared, adapted mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. Coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

Query Match 100.0%; Score 101; DB 9; Length 519;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTTAGAAAAATAACAATAG 60
 |||||
Db 416 AGGGTTATTGTCATGAGCGGATACATATTTGAATGTTAGAAAAATAACAATAG 357
 |||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
 |||||
Db 356 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 316
 |||||

Search completed: July 14, 2005, 23:23:20
 Job time : 962.667 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:39:07 ; Search time 749.127 Seconds
(without alignments)

6468.225 Million cell updates/sec

Title: US-09-482-682-65_COPY_1_100

Perfect score: 100

Sequence: 1 ctgtccctcgtgtgtgtt.....caattgatgaagaatctgc 100

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.on.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100	100.0	562	6	AX643583 Sequence
2	100	100.0	633	14	ALRPROLTB
3	100	100.0	648	6	AX175190 Sequence
4	100	100.0	648	6	AX175195 Sequence
5	100	100.0	1070	6	A85308 Sequence 6
6	100	100.0	1070	6	BD107647 FIV vacci
7	100	100.0	2245	6	AX643582 Sequence
8	100	100.0	2426	6	AX044426 Sequence
9	100	100.0	2427	6	AX044425 Sequence
10	100	100.0	3557	12	SYNRSV3MV
11	100	100.0	3840	12	EV132038
12	100	100.0	3853	6	AR098190 Sequence
13	100	100.0	3853	6	AR207832 Sequence
14	100	100.0	3853	6	BD009729 Tissue sp
15	100	100.0	3925	6	A60213 Sequence 9
16	100	100.0	3925	6	ARI22289 Sequence
17	100	100.0	3986	12	PCDNA32EO
18	100	100.0	4026	6	AR098191 Sequence
19	100	100.0	4026	6	AR207833 Sequence

20	100	100.0	4026	6	BD009730
21	100	100.0	4059	6	AR071324 Sequence
22	100	100.0	4249	6	AR098192 Sequence
23	100	100.0	4249	6	AR207834 Sequence
24	100	100.0	4249	6	BD009731 Tissue sp
25	100	100.0	4341	6	A38214 Sequence 58
26	100	100.0	4341	6	AX286570 Sequence
27	100	100.0	4457	6	AX743954 Sequence
28	100	100.0	4525	6	AR062871 Sequence
29	100	100.0	4597	6	AX060344 Sequence
30	100	100.0	4839	12	SYNRSV5GPT
31	100	100.0	4840	6	AX133940 Sequence
32	100	100.0	4965	6	AR071323 Sequence
33	100	100.0	5053	6	BD238492 Expressio
34	100	100.0	5070	6	AX234391 Sequence
35	100	100.0	5082	6	A91754 Sequence 10
36	100	100.0	5082	6	BD085110 Vertebrat
37	100	100.0	5108	12	SYNRSV5NEO
38	100	100.0	5162	6	AX951626 Sequence
39	100	100.0	5257	12	CVU89673
40	100	100.0	5432	6	BD234590 Screening
41	100	100.0	5432	6	AX026821 Sequence
42	100	100.0	5446	6	BD195386 Compositi
43	100	100.0	5446	6	AX319694 Sequence
44	100	100.0	5564	12	SYNTRC
45	100	100.0	5618	6	A44171 Sequence 1

ALIGNMENTS

RESULT 1
AX643583
LOCUS AX643583 562 bp DNA linear PAT 24-FEB-2003
DEFINITION Sequence 2 from Patent WO02099100.
ACCESSION AX643583
VERSION AX643583.1 GI:28551383
KEYWORDS
SOURCE Mus sp.
ORGANISM Mus sp.
REFERENCE 1
AUTHORS Al-Rubeai, M. and Shuttleworth, J.
TITLE Method of production of a protein in cells which inducibly express the cell cycle inhibitor protein, p21
JOURNAL Patent: WO 02099100-A 2 12-DEC-2002;
FEATURES Lonza Biologics plc (GB)
Location/Qualifiers
source 1..562
/organism="Mus sp."
/mol_type="unassigned DNA"
/db_xref="taxon:10095"
/note="Rous Sarcoma Virus LTR promoter"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 562;
Best Local Similarity 100.0%; Pred. No. 2.6e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTGCTCCCTGTTGTGTGGAGGTGCTGCTGAGTAGTGCAGCGACGACAAATTAAGCTACA 60
Db 46 CTGCTCCCTGTTGTGTGTGGAGGTGCTGCTGAGTAGTGCAGCGACGACAAATTAAGCTACA 105
Qy 61 ACAAGGCAAGCTTGACCGCAATTCATGAAGAATCTGC 100.
Db 106 ACAAGGCAAGCTTGACCGCAATTCATGAAGAATCTGC 145

RESULT 2
ALRPROLTB
LOCUS ALRPROLTB 633 bp ss-RNA linear VRL 28-APR-1993
DEFINITION Rous sarcoma virus (Schmidt-Ruppin), proviral, 3' LTR on 21S mRNA.

ACCESSION J02025 J02022
VERSION J02025.1 GI:210255
KEYWORDS c-myc proto-oncogene; long terminal repeat (LTR); src oncogene.
SOURCE Rous sarcoma virus
ORGANISM Rous sarcoma virus
REFERENCE 1 (sites)
AUTHORS Yamamoto, T., de Crombrughe, B. and Pastan, I.
TITLE Identification of a functional promoter in the long terminal repeat of Rous sarcoma virus
JOURNAL Cell 22 (3), 787-797 (1980)
MEDLINE 81112147
PUBMED 6257399
REFERENCE 2 (bases 1 to 633)
AUTHORS Yamamoto, T., Tyagi, J.S., Fagan, J.B., Jay, G., deCrombrughe, B. and Pastan, I.
TITLE Molecular mechanism for the capture and excision of the transforming gene of avian sarcoma virus as suggested by analysis of recombinant clones
JOURNAL J. Virol. 35 (2), 436-443 (1980)
MEDLINE 81072438
PUBMED 6255184
REFERENCE 3 (bases 319 to 633)
AUTHORS Yamamoto, T., Jay, G. and Pastan, I.
TITLE Unusual features in the nucleotide sequence of a cDNA clone derived from the common region of avian sarcoma virus messenger RNA
Proc. Natl. Acad. Sci. U.S.A. 77 (1), 176-180 (1980)
JOURNAL 80145590
MEDLINE 6244542
PUBMED
COMMENT Original source text: Rous sarcoma virus (Schmidt-Ruppin strain, subgroup D) provirus, cDNA to 21S mRNA from infected chicken embryonic fibroblasts, clone pSR1.
[1] sites; mRNA start.
Original figure in [2] included 24 'g's on 5' end and 16 'c's on 3' end that were cDNA synthesis artifacts.
[2] also sequenced a defective clone, pSR2, with the src gene deleted (see separate entry).
[1] demonstrated the mRNA transcription initiation site shown in the Sites table using pSR1 as a template. However, this is the 3' LTR, and the functional mRNA start site would be assumed to be on the 5' LTR at the homologous site.
FEATURES
source Location/Qualifiers
1..633
/organism="Rous sarcoma virus"
/mol_type="genomic RNA"
/db_xref="taxon:11886"
misc_RNA <1..517
/note="viral genomic RNA"
LTR 211..5633
/note="3' LTR"
mRNA 517..5633
/note="in vitro mRNA [1]; see comment"
repeat_region 517..536
/note="terminally redundant repeat"
ORIGIN 20 bp upstream of PstI site.
Query Match 100.0%; Score 100; DB 14; Length 633;
Best Local Similarity 100.0%; Pred. No. 2.6e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTGCTCCCTGCTGTGTGGAGTCTGCTAGTAGTCGCGAGCAAAATTTAAGCTACA 60
Db 28 CTGCTCCCTGCTGTGTGTGGAGTCTGCTAGTAGTCGCGAGCAAAATTTAAGCTACA 87
Qy 61 ACAAGGCAAGGCTTGACCGACAATTCGATGAAGAATCTGC 100
Db 88 ACAAGGCAAGGCTTGACCGACAATTCGATGAAGAATCTGC 127
RESULT 3
AX175190
LOCUS Sequence 1 from Patent WO014244. 648 bp DNA linear PAT 03-JUL-2001
DEFINITION

AX175190
VERSION AX175190.1 GI:14598581
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Rivera, V., Zoltick, P. and Wilson, J.M.
TITLE Methods for expression of genes in primates
JOURNAL Patent: WO 014244-A 1 14-JUN-2001;
ARIAD GENE THERAPEUTICS, INC. (US); THE UNIVERSITY OF PENNSYLVANIA (US)
FEATURES
source Location/Qualifiers
1..648
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="vector/RSV promoter/vector"
ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 648;
Best Local Similarity 100.0%; Pred. No. 2.6e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTGCTCCCTGCTGTGTGTGGAGTCTGCTAGTAGTCGCGAGCAAAATTTAAGCTACA 60
Db 90 CTGCTCCCTGCTGTGTGTGGAGTCTGCTAGTAGTCGCGAGCAAAATTTAAGCTACA 149
Qy 61 ACAAGGCAAGGCTTGACCGACAATTCGATGAAGAATCTGC 100
Db 150 ACAAGGCAAGGCTTGACCGACAATTCGATGAAGAATCTGC 189
RESULT 4
AX175195
LOCUS Sequence 6 from Patent WO014244. 648 bp DNA linear PAT 03-JUL-2001
DEFINITION
ACCESSION AX175195
VERSION AX175195.1 GI:14598586
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Rivera, V., Zoltick, P. and Wilson, J.M.
TITLE Methods for expression of genes in primates
JOURNAL Patent: WO 014244-A 6 14-JUN-2001;
ARIAD GENE THERAPEUTICS, INC. (US); THE UNIVERSITY OF PENNSYLVANIA (US)
FEATURES
source Location/Qualifiers
1..648
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="MLuI/RSV promoter/BglI"
ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 648;
Best Local Similarity 100.0%; Pred. No. 2.6e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTGCTCCCTGCTGTGTGTGGAGTCTGCTAGTAGTCGCGAGCAAAATTTAAGCTACA 60
Db 90 CTGCTCCCTGCTGTGTGTGGAGTCTGCTAGTAGTCGCGAGCAAAATTTAAGCTACA 149
Qy 61 ACAAGGCAAGGCTTGACCGACAATTCGATGAAGAATCTGC 100
Db 150 ACAAGGCAAGGCTTGACCGACAATTCGATGAAGAATCTGC 189
RESULT 5
AX175190
LOCUS Sequence 1 from Patent WO014244. 1070 bp DNA linear PAT 21-JAN-2000
DEFINITION

```

DEFINITION Sequence 6 from Patent WO9840493.
ACCESSION A85308
VERSION A85308.1 GI:6733916
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
    source
        1..1070
        /organism="unidentified"
        /mol_type="unassigned DNA"
        /db_xref="taxon:32644"
ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 1070;
Best Local Similarity 100.0%; Pred. No. 2.8e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTGTGTGAGTGCCTGAGTAGTCGGCGAGCAAAATTTAAGCTACA 60
Db 77 CTGCTCCCTGCTGTGTGAGTGCCTGAGTAGTCGGCGAGCAAAATTTAAGCTACA 136

Qy 61 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 100
Db 137 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 176

RESULT 7
AX643582
LOCUS AX643582 2245 bp DNA linear PAT 24-FEB-2003
DEFINITION Sequence 1 from Patent WO2099100.
ACCESSION AX643582
VERSION AX643582.1 GI:28551382
KEYWORDS
SOURCE
ORGANISM Mus sp.
REFERENCE
AUTHORS Al-Rubeai, M. and Shuttlesworth, J.
TITLE Method of production of a protein in cells which inducibly express the cell cycle inhibitor protein, p21
JOURNAL Patent: WO 02099100-A 1 12-DEC-2002; Lonza Biologics plc (GB)
FEATURES
    source
        1..2245
        /organism="Mus sp."
        /mol_type="unassigned DNA"
        /db_xref="taxon:10095"
        /note="RSV-LTR promoter + intron + p21 cds + Tkpoly(A) LacSwitch II expression construct"
ORIGIN

Query Match 100.0%; Score 100; DB 6; Length 2245;
Best Local Similarity 100.0%; Pred. No. 3e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTGTGTGAGTGCCTGAGTAGTCGGCGAGCAAAATTTAAGCTACA 60
Db 46 CTGCTCCCTGCTGTGTGAGTGCCTGAGTAGTCGGCGAGCAAAATTTAAGCTACA 105

Qy 61 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 100
Db 106 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 145

RESULT 8
AX044426
LOCUS AX044426 2426 bp DNA linear PAT 24-NOV-2000
DEFINITION Sequence 18 from Patent WO0066752.
ACCESSION AX044426
VERSION AX044426.1 GI:11343299
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Castro, M.G., Emery, S.C. and Lowenstein, P.R.
TITLE Chemical compounds
JOURNAL Patent: WO 0066752-A 18 09-NOV-2000; Astrazeneca AB (SE); THE VICTORIA UNIVERSITY OF MANCHESTER (GB)
FEATURES
    source
        1..2426
        /organism="synthetic construct"
        /mol_type="unassigned DNA"
        /db_xref="taxon:32630"
        /note="CPG2 with last exon of Thy-1 fused at 3' end"
ORIGIN

Query Match 100.0%; Score 100; DB 6; Length 1070;
Best Local Similarity 100.0%; Pred. No. 2.8e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTGTGTGAGTGCCTGAGTAGTCGGCGAGCAAAATTTAAGCTACA 60
Db 77 CTGCTCCCTGCTGTGTGAGTGCCTGAGTAGTCGGCGAGCAAAATTTAAGCTACA 136

Qy 61 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 100
Db 137 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 176

RESULT 6
BD107647
LOCUS BD107647 1070 bp DNA linear PAT 18-SEP-2002
DEFINITION FIV vaccine.
ACCESSION BD107647
VERSION BD107647.1 GI:23202465
KEYWORDS JP 2002501369-A/6.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE
AUTHORS Neil, J.C., Rigby, M.A. and Jarrett, J.O.
TITLE FIV vaccine
JOURNAL Patent: JP 2002501369-A 6 15-JAN-2002; THE UNIVERSITY COURT OF THE UNIVERSITY OF GLASGOW
COMMENT OS Artificial Sequence
        PN JP 2002501369-A/6
        PD 15-JAN-2002
        PF 10-MAR-1998 JP 1998539351
        PI JAMES CHARLES NEIL, MARK ALAN RIGBY, JAMES OSWALD JARRETT PC
        CC C12N15/49, A61K31/70, A61K48/00
        CC CMV PROMOTER FROM pCDNA3 (a Bgl II - Kpn I restriction fragment)
        CC SST I - SST I FRAGMENT IN PLASMID CMV DEL. RT CC FIV GENOME
        FROM THE t-RNA PRIMER BINDING SITE TO THE VIRAL Sst IS
        CC
        FT Key Location/Qualifiers
        FT source 1..1070
        FT /organism="Artificial Sequence".
        FT Location/Qualifiers
        1..1070
        /organism="synthetic construct"
        /mol_type="genomic DNA"
        /db_xref="taxon:32630"
FEATURES
    source
ORIGIN

Query Match 100.0%; Score 100; DB 6; Length 1070;
Best Local Similarity 100.0%; Pred. No. 2.8e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTGTGTGAGTGCCTGAGTAGTCGGCGAGCAAAATTTAAGCTACA 60
Db 77 CTGCTCCCTGCTGTGTGAGTGCCTGAGTAGTCGGCGAGCAAAATTTAAGCTACA 136

Qy 61 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 100
Db 137 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 176

```

Query Match 100.0%; Score 100; DB 6; Length 2426;
 Best Local Similarity 100.0%; Pred. No. 3.1e-25;
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTTGTGTGGAGTCCCTGAGTAGTGGCGAGCAAAATTTAAAGCTACA 60
 Db 69 CTGCTCCCTGCTTGTGTGTGGAGTCCCTGAGTAGTGGCGAGCAAAATTTAAAGCTACA 128

Qy 61 ACAAGGCAAGCTTGACCGACAATTGCGATGAAGAATCTGC 100
 Db 129 ACAAGGCAAGCTTGACCGACAATTGCGATGAAGAATCTGC 168

RESULT 9
 LOCUS AX044425 2427 bp DNA linear PAT 24-NOV-2000
 DEFINITION Sequence 17 from Patent WO0066752.
 ACCESSION AX044425
 VERSION AX044425.1 GI:11343298
 KEYWORDS synthetic construct
 SOURCE synthetic construct
 ORGANISM other sequences; artificial sequences.

REFERENCE 1
 AUTHORS Castro, M.G., Emery, S.C. and Lowenstein, P.R.
 TITLE Chemical compounds
 JOURNAL Patent: WO 0066752-A 17 09-NOV-2000;
 AstraZeneca AB (SE); THE VICTORIA UNIVERSITY OF MANCHESTER (GB)

FEATURES
 Location/Qualifiers
 source 1..2427
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"
 /note="CPG2 mutant with last exon of Thy-1 fused at 3' end"

ORIGIN
 Query Match 100.0%; Score 100; DB 6; Length 2427;
 Best Local Similarity 100.0%; Pred. No. 3.1e-25;
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTTGTGTGGAGTCCCTGAGTAGTGGCGAGCAAAATTTAAAGCTACA 60
 Db 70 CTGCTCCCTGCTTGTGTGTGGAGTCCCTGAGTAGTGGCGAGCAAAATTTAAAGCTACA 129

Qy 61 ACAAGGCAAGCTTGACCGACAATTGCGATGAAGAATCTGC 100
 Db 130 ACAAGGCAAGCTTGACCGACAATTGCGATGAAGAATCTGC 169

RESULT 10
 SYNRSV3MV
 LOCUS SYNRSV3MV 3557 bp DNA circular SYN 27-APR-1993
 DEFINITION Cloning vector RSV3.
 ACCESSION M83240
 VERSION M83240.1 GI:209303
 KEYWORDS cDNA expression vector.
 SOURCE unidentified cloning vector
 ORGANISM unidentified cloning vector
 other sequences; artificial sequences; vectors.

REFERENCE 1
 AUTHORS Messing, J.
 TITLE New M13 vectors for cloning
 JOURNAL Meth. Enzymol. 101, 20-78 (1983)
 MEDLINE 83296918
 PUBMED 6310323

REFERENCE 2
 (sites)
 Gorman, C., Padmanabhan, R. and Howard, B.H.
 TITLE High efficiency DNA-mediated transformation of primate cells
 JOURNAL Science 221 (4610), 551-553 (1983)
 MEDLINE 83249156
 PUBMED 6306768

REFERENCE 3
 (bases 1 to 3557)

AUTHORS Jacobson, S., Sekaly, R.P., Jacobson, C.L., McFarland, H.F. and Long, F.O.
 TITLE HLA class II-restricted presentation of cytoplasmic measles virus antigens to cytotoxic T cells
 JOURNAL J. Virol. 63 (4), 1756-1762 (1989)
 MEDLINE 89178863
 PUBMED 2784598

COMMENT Original source text: Cloning vector DNA.
 FEATURES
 Location/Qualifiers
 source 1..3557
 /organism="unidentified cloning vector"
 /mol_type="genomic DNA"
 /db_xref="taxon:45196"
 misc_feature 1..29
 /function="polylinker"
 /evidence="experimental"
 misc_feature 912..3029
 /function="ampicillin-resistance, replication origin"
 /evidence="experimental"
 enhancer 3030..3557
 /standard_name="5' LTR of Rous Sarcoma Virus"
 /citation=[2]
 /evidence="experimental"

ORIGIN
 Query Match 100.0%; Score 100; DB 12; Length 3557;
 Best Local Similarity 100.0%; Pred. No. 3.2e-25;
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTTGTGTGGAGTCCCTGAGTAGTGGCGAGCAAAATTTAAAGCTACA 60
 Db 3030 CTGCTCCCTGCTTGTGTGGAGTCCCTGAGTAGTGGCGAGCAAAATTTAAAGCTACA 3089

Qy 61 ACAAGGCAAGCTTGACCGACAATTGCGATGAAGAATCTGC 100
 Db 3090 ACAAGGCAAGCTTGACCGACAATTGCGATGAAGAATCTGC 3129

RESULT 11
 EVE132038
 LOCUS EVE132038 3840 bp RNA circular SYN 28-JUL-1999
 DEFINITION Expression vector pCDPT.
 ACCESSION AJ132038
 VERSION AJ132038.1 GI:5640088
 KEYWORDS AMP gene; beta lactamase; ColE1 origin of replication; multiple cloning site; SP6 promoter; SV40 origin of replication; T7 promoter; xanthine-guanine phosphoribosyl transferase; xanthine-guanine phosphoribosyl transferase gene.

SOURCE Expression vector pCDPT
 ORGANISM Expression vector pCDPT
 other sequences; artificial sequences; vectors.

REFERENCE 1
 Zeng, B.J.
 AUTHORS Mammalian Expression Vector for with fuse Xanthine-guanine phosphoribosyl transferase Tag
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 3840)
 Zeng, B.J.
 AUTHORS Direct Submission
 JOURNAL Submitted (27-FEB-1999) Zeng B.J., Gene Engineering Center, Institute of Microbiology, Zhongguancun, Beijing, Beijing 100080, CHINA

FEATURES
 Location/Qualifiers
 source 1..3840
 /organism="Expression vector pCDPT"
 /mol_type="Other RNA"
 /db_xref="taxon:90749"
 promoter 209..863
 /note="CMV"
 promoter 864..882
 /note="T7"
 misc_feature 882..984
 /note="Multiple cloning site"


```
CDS
HindIII, BamHI, BstXI, EcoRI, NotI, XhoI
929..1387
/codon_start=1
/product="Xanthine-guanine phosphoribosyl transferase"
/protein_id="CAB51567.1"
/db_xref="GI:5640089"
/translation="MSEKIIVTMDLQIHARKLASRLMPSEQWKGIIAVSRGGLVPGA
LLARELGHVDTVCISYDHNQRELKVLKRAEGDGEFVIDDLVDTGTAVAIKE
MYPKAHFVTIIFAKPAGRLVDVYVDIPQDTWIEQPMGVMVFVPPISGR"
1649..1863
/feature="BCH"
2450..2775
/feature="SP6"
2644..2729
/feature="SV40"
complement(2844..3704)
/feature="amp"
complement(2844..3704)
/feature="amp"
/codon_start=1
/product="beta-lactamase"
/protein_id="CAB51568.1"
/db_xref="GI:5640090"
/translation="MSIQHFRVALIPFAAFCLPVFAHPETLVKVKDAEDQLGARVY
IELDLSGKILESFRPEFPMSTFKVLLCGAVLSRIDAGEQLRIHYSQNDLVE
YSPVTEKHLTDGMTRELCSAAITMSDNTAANLLLTIGGPKELTAFIHNMGDVTRL
DRWPELNEATPNDERTTPMVAMATLRKLLTGELTLASRQQLIDWMEADKVAQPL
LRSLPAGWFIADKSGAGERSGIIAALGPDGKPSRIVVIYTGTSQATWDERNRQIA
EIGASLKHWH"
3632..3840
/feature="ColE1"

rep_origin
Query Match 100.0%; Score 100; DB 12; Length 3840;
Best Local Similarity 100.0%; Pred. No. 3.3e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTTGTGTGGAGTCCCTGAGTAGTCGCGGAGCAAAATTTAAGCTACA 60
Db 6 CTGCTCCCTGCTTGTGTGGAGTCCCTGAGTAGTCGCGGAGCAAAATTTAAGCTACA 65

Qy 61 ACAAGGCAAGCTTGACCGACAATTCGATGAAGAATCTGC 100
Db 66 ACAAGGCAAGCTTGACCGACAATTCGATGAAGAATCTGC 105

RESULT 12
AR098190 LOCUS 3853 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 5 from patent US 6074850.
ACCESSION AR098190
VERSION AR098190.1 GI:12807447
KEYWORDS Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3853)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion polypeptides
JOURNAL Patent: US 6074850-A 5 13-JUN-2000;
FEATURES Location/Qualifiers
source 1..3853
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 3853;
Best Local Similarity 100.0%; Pred. No. 3.3e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTTGTGTGGAGTCCCTGAGTAGTCGCGGAGCAAAATTTAAGCTACA 60
Db 81 CTGCTCCCTGCTTGTGTGGAGTCCCTGAGTAGTCGCGGAGCAAAATTTAAGCTACA 140

CDS
HindIII, BamHI, BstXI, EcoRI, NotI, XhoI
929..1387
/codon_start=1
/product="Xanthine-guanine phosphoribosyl transferase"
/protein_id="CAB51567.1"
/db_xref="GI:5640089"
/translation="MSEKIIVTMDLQIHARKLASRLMPSEQWKGIIAVSRGGLVPGA
LLARELGHVDTVCISYDHNQRELKVLKRAEGDGEFVIDDLVDTGTAVAIKE
MYPKAHFVTIIFAKPAGRLVDVYVDIPQDTWIEQPMGVMVFVPPISGR"
1649..1863
/feature="BCH"
2450..2775
/feature="SP6"
2644..2729
/feature="SV40"
complement(2844..3704)
/feature="amp"
complement(2844..3704)
/feature="amp"
/codon_start=1
/product="beta-lactamase"
/protein_id="CAB51568.1"
/db_xref="GI:5640090"
/translation="MSIQHFRVALIPFAAFCLPVFAHPETLVKVKDAEDQLGARVY
IELDLSGKILESFRPEFPMSTFKVLLCGAVLSRIDAGEQLRIHYSQNDLVE
YSPVTEKHLTDGMTRELCSAAITMSDNTAANLLLTIGGPKELTAFIHNMGDVTRL
DRWPELNEATPNDERTTPMVAMATLRKLLTGELTLASRQQLIDWMEADKVAQPL
LRSLPAGWFIADKSGAGERSGIIAALGPDGKPSRIVVIYTGTSQATWDERNRQIA
EIGASLKHWH"
3632..3840
/feature="ColE1"

rep_origin
Query Match 100.0%; Score 100; DB 12; Length 3840;
Best Local Similarity 100.0%; Pred. No. 3.3e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTTGTGTGGAGTCCCTGAGTAGTCGCGGAGCAAAATTTAAGCTACA 60
Db 6 CTGCTCCCTGCTTGTGTGGAGTCCCTGAGTAGTCGCGGAGCAAAATTTAAGCTACA 65

Qy 61 ACAAGGCAAGCTTGACCGACAATTCGATGAAGAATCTGC 100
Db 66 ACAAGGCAAGCTTGACCGACAATTCGATGAAGAATCTGC 105

RESULT 12
AR098190 LOCUS 3853 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 5 from patent US 6074850.
ACCESSION AR098190
VERSION AR098190.1 GI:12807447
KEYWORDS Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3853)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion polypeptides
JOURNAL Patent: US 6074850-A 5 13-JUN-2000;
FEATURES Location/Qualifiers
source 1..3853
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 3853;
Best Local Similarity 100.0%; Pred. No. 3.3e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTTGTGTGGAGTCCCTGAGTAGTCGCGGAGCAAAATTTAAGCTACA 60
Db 81 CTGCTCCCTGCTTGTGTGGAGTCCCTGAGTAGTCGCGGAGCAAAATTTAAGCTACA 140
```

```
Qy 61 ACAAGGCAAGCTTGACCGACAATTCGATGAAGAATCTGC 100
Db 141 ACAAGGCAAGCTTGACCGACAATTCGATGAAGAATCTGC 180

RESULT 13
AR207832 LOCUS 3853 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 5 from patent US 6379927.
ACCESSION AR207832
VERSION AR207832.1 GI:21507688
KEYWORDS Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3853)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion proteins
JOURNAL Patent: US 6379927-A 5 30-APR-2002;
FEATURES Location/Qualifiers
source 1..3853
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 3853;
Best Local Similarity 100.0%; Pred. No. 3.3e-25;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTTGTGTGGAGTCCCTGAGTAGTCGCGGAGCAAAATTTAAGCTACA 60
Db 81 CTGCTCCCTGCTTGTGTGGAGTCCCTGAGTAGTCGCGGAGCAAAATTTAAGCTACA 140

Qy 61 ACAAGGCAAGCTTGACCGACAATTCGATGAAGAATCTGC 100
Db 141 ACAAGGCAAGCTTGACCGACAATTCGATGAAGAATCTGC 180

RESULT 14
BD009729 LOCUS 3853 bp DNA linear PAT 31-JAN-2002
DEFINITION Tissue specific expression of retinoblastoma protein.
ACCESSION BD009729
VERSION BD009729.1 GI:18638102
KEYWORDS JP 2001503638-A/3.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 3853)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Tissue specific expression of retinoblastoma protein
JOURNAL Patent: JP 2001503638-A 3 21-MAR-2001;
COMMENT CANJI INC
OS Unidentified
PN JP 2001503638-A/3
PD 21-MAR-2001
PF 13-NOV-1997 JP 1998522958
PR 15-NOV-1996 US 08/751517,14-FEB-1997 US 08/801092 PI
DOUGLAS ANTELMAN,RICHARD J GREGORY,KENNETH N WILLS PC
C07H21/04,C07K5/00,A61K38/00,A61K35/12
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT CDS Location/Qualifiers
source 1..3853
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 3853;
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:35:42 ; Search time 140.988 Seconds

(without alignments)
4198.742 Million cell updates/sec

Title: US-09-482-682-65_COPY_1_100

Perfect score: 100

Sequence: 1 ctgtcctcgtgtgtgtgtt.....caattgcatgaagaatctgc 100

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04.*

1: Geneseqn1980s.*

2: Geneseqn1990s.*

3: Geneseqn2000s.*

4: Geneseqn2001as.*

5: Geneseqn2001bs.*

6: Geneseqn2002as.*

7: Geneseqn2002bs.*

8: Geneseqn2003as.*

9: Geneseqn2003bs.*

10: Geneseqn2003cs.*

11: Geneseqn2003ds.*

12: Geneseqn2004as.*

13: Geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100	100.0	562	8	ABZ23250 Nucleotid
2	100	100.0	648	4	Aah43951 Rous sarc
3	100	100.0	1070	2	Aav58058 Plasmid C
4	100	100.0	1506	12	Adm41035 Fungus nu
5	100	100.0	1600	2	Adh11349 Vertebrat
6	100	100.0	1782	12	Adm41037 Cytomegal
7	100	100.0	2241	12	Adm41034 Human nuc
8	100	100.0	2245	8	ABZ23249 Lac repre
9	100	100.0	2294	12	Adm41036 Cytomegal
10	100	100.0	2426	4	Aad02037 Plasmid p
11	100	100.0	2427	4	Aad02036 Plasmid p
12	100	100.0	3400	2	Aat62937 3F4 human
13	100	100.0	3400	2	Aat62932 2A2 human
14	100	100.0	3853	2	Aav40006 Plasmid p
15	100	100.0	3925	2	Aat90695 Plasmid C
16	100	100.0	4026	2	Aav40007 Plasmid p
17	100	100.0	4059	2	Aaq75974 pHLA-B7 e
18	100	100.0	4249	2	Aav63466 Plasmid p
19	100	100.0	4341	2	Aaq62391 Vector pV
20	100	100.0	4341	6	Aas17704 Vector pV

21	100	100.0	4341	6	ABN83143	Abn83143 Plasmid p
22	100	100.0	4457	10	ADD35599	Add35599 Bicistron
23	100	100.0	4525	2	AAV69746	Aav69746 Nucleotid
24	100	100.0	4597	4	AAF24901	Aaf24901 Nucleotid
25	100	100.0	4825	13	ADR12380	Adr12380 Vector pM
26	100	100.0	4840	4	AAF83146	Aaf83146 Complete
27	100	100.0	4965	2	AAQ75973	Aaq75973 pHLA-B7/b
28	100	100.0	5015	10	ADB33528	Adb33528 Expressio
29	100	100.0	5053	3	AAZ38633	Aaz38633 pEP2 expr
30	100	100.0	5070	4	AAS12839	Aas12839 DNA sequ
31	100	100.0	5082	2	ADH11417	Adh11417 plasmid p
32	100	100.0	5162	10	ADF10526	Adf10526 Plasmid p
33	100	100.0	5162	10	ACC44637	Acc44637 Murine rD
34	100	100.0	5172	13	ADS75099	Ads75099 Plasmid p
35	100	100.0	5173	6	ABK88869	Abk88869 Topoisome
36	100	100.0	5173	12	ADR83792	Adr83792 Plasmid p
37	100	100.0	5173	12	ADO06721	Ado06721 Recombina
38	100	100.0	5192	10	ACC44692	Acc44692 Plasmid p
39	100	100.0	5250	2	AAT62933	Aat62933 2A2 human
40	100	100.0	5271	10	ABV77540	Abv77540 Plasmid p
41	100	100.0	5283	10	ABV77538	Abv77538 Plasmid p
42	100	100.0	5292	10	ABV77547	Abv77547 Plasmid p
43	100	100.0	5293	10	ABV77548	Abv77548 Plasmid p
44	100	100.0	5293	10	ABV77549	Abv77549 Plasmid p
45	100	100.0	5300	2	AAT62938	Aat62938 3F4 human

ALIGNMENTS

RESULT 1

ABZ23250

ID ABZ23250 standard; DNA; 562 BP.

AC ABZ23250;

XX 24-MAR-2003 (first entry)

XX Nucleotide sequence of the Rous sarcoma virus (RSV)-LTR promoter.

XX p21; RSV; LTR promoter; cell cycle inhibitor protein; protein production;
XX anchorage-independent producer cell line; ss.
XX Rous sarcoma virus.
XX WO200299100-A2.
XX 12-DEC-2002.
XX 03-JUN-2002; 2002WO-EP006054.
XX 01-JUN-2001; 2001GB-00013318.
XX (LONZ) LONZA BIOLOGICS PLC.
XX Al-Rubeai M, Shuttleworth J;
XX WPI; 2003-148669/14.
XX Producing recombinant protein, particularly for maximizing or enhancing
XX e.g. therapeutic protein production, by co-expressing protein with
XX recombinant cell cycle inhibitor protein (p21) in producer cell line.
XX Disclosure; Page 32-33; 33pp; English.
XX The present sequence represents the Rous sarcoma virus (RSV)-LTR
XX promoter. The present sequence is used to produce vectors for use in the
XX method of the invention. The specification describes a method for
XX producing a protein, preferably a recombinant protein, in a mammalian
XX anchorage-independent producer cell line. The method comprises co-
XX expressing with the protein in the producer cell line a recombinant cell
XX cycle inhibitor protein (preferably p21). The method is useful for
XX producing a recombinant protein in a producer cell line. This is

CC particularly useful for maximizing or enhancing the production of e.g.
XX therapeutic proteins at an industrial scale
SQ Sequence 562 BP; 143 A; 109 C; 163 G; 147 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 8; Length 562;
Best Local Similarity 100.0%; Pred. No. 8.2e-28;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTTGTGTTGGAGTGCCTGAGTAGTGCAGGACAAATTTAAGCTACA 60
Db 46 CTGCTCCCTGCTTGTGTTGGAGTGCCTGAGTAGTGCAGGACAAATTTAAGCTACA 105

Qy 61 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 100
Db 106 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 145

RESULT 2
AAH43951
ID AAH43951 standard; DNA; 648 BP.
AC AAH43951;
DT 06-SEP-2001 (first entry)
DE Rous sarcoma virus promoter nucleotide sequence SEQ ID NO:1.
XX Rous sarcoma virus; promoter; enhancer; RSV; primate; gene expression;
KW transgene; genetic engineering; gene therapy; immunisation; ds.
XX Rous sarcoma virus.
XX WO200142444-A2.
XX 14-JUN-2001.
XX 08-DEC-2000; 2000WO-US033256.
XX 10-DEC-1999; 99US-0170019P.
XX (ARIA-) ARIAD GENE THERAPEUTICS INC.
XX (UYPE-) UNIV PENNSYLVANIA.
XX Rivera V, Zoltick P, Wilson JM;
XX WPI; 2001-381673/40.
XX Genetically engineering a primate for expression of a desired gene,
XX comprises introducing into the primate a transgene comprising Rous
XX Sarcoma Virus (RSV) promoter and a nucleic acid sequence heterologous to
XX RSV promoter.
XX Claim 7; Page 44; 64pp; English.
XX
XX The present invention describes a method for genetically engineering a
XX primate for expression of a desired gene comprising introducing into the
XX primate a transgene comprising an Rous Sarcoma Virus (RSV) promoter and a
XX nucleic acid sequence heterologous to RSV promoter. Also described is a
XX primate cell (i) containing and capable of expressing a transgene
XX comprising an RSV promoter operably linked to a recombinant nucleic acid
XX encoding one or more fusion proteins, where the fusion proteins bind to a
XX ligand and in the presence of the ligand modulate(s) the expression of
XX a target gene. The method can be used for high level expression of
XX genes in primates or for engineering primate cells. It is useful for
XX increasing the efficacy of many gene therapy strategies, and for
XX like ribozymes, antisense RNA, and dominant negative proteins, that act
XX either stoichiometrically, or by competition. The method increases the
XX efficacy of many gene therapy strategies by substantially elevating the
XX expression of an exogenous therapeutic gene, and allowing expression to
XX reach therapeutically effective levels. The present sequence represents a
XX specifically claimed RSV enhancer/promoter nucleotide sequence from the

CC present invention
XX
SQ Sequence 648 BP; 163 A; 135 C; 179 G; 171 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 4; Length 648;
Best Local Similarity 100.0%; Pred. No. 8.6e-28;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTTGTGTTGGAGTGCCTGAGTAGTGCAGGACAAATTTAAGCTACA 60
Db 90 CTGCTCCCTGCTTGTGTTGGAGTGCCTGAGTAGTGCAGGACAAATTTAAGCTACA 149

Qy 61 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 100
Db 150 ACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGC 189

RESULT 3
AAV58058
ID AAV58058 standard; DNA; 1070 BP.
XX
AC AAV58058;
XX
XX 27-AUG-2003 (revised)
DT 11-JAN-1999 (first entry)
DE Plasmid CMV-delRT SstI fragment.
XX
XX FIV; FIPV; vaccine; reverse transcriptase; diagnosis; therapy; CMV-delRT;
KW promoter; cat; ss.
XX Human cytomegalovirus.
OS feline immunodeficiency virus.
OS Chimeric.
XX
XX Key Location/Qualifiers
FT promoter 8..896
FT /tag= a
FT /note= "CMV promoter fragment from pCDNA3 (BgIII-KpnII)"
FT provirus 918..1070
FT /tag= b
FT /note= "FIV sequences from primer binding site to SstI
FT site"
XX
XX WO9840493-A1.
XX
XX 17-SEP-1998.
XX
XX 10-MAR-1998; 98WO-GB0000715.
XX
XX 11-MAR-1997; 97GB-00004977.
XX
XX (UNIU) UNIV GLASGOW.
XX
XX Neil JC, Rigby MA, Jarrett JO;
XX WPI; 1998-520813/44.
XX
XX Protecting, e.g. cats, against feline immunodeficiency virus - by using
XX vaccine comprising FIV pol gene containing deletion and/or insertion in
XX reverse transcriptase domain.
XX
XX Example 3; Fig 4; 66pp; English.
XX
XX This is the nucleotide sequence of a SstI fragment of plasmid CMV-delRT,
XX in which the immediate-early promoter of human cytomegalovirus replaces
XX the 5' long terminal repeat region of feline immunodeficiency virus (FIV)
XX clone F14-delRT (see AAV58053). FIV sequences downstream of the SstI site
XX are identical to those in F14-delRT. Use of the CMV promoter was designed
XX to enhance expression of FIV antigens, and to reduce the risk of
XX reversion to a replicating provirus, in tissues after inoculation of DNA.
XX Vaccine formulations for FIV-related diseases include a defective feline
XX immunodeficiency proviral (FIPV) polynucleotide comprising an in-frame

CC active in the signal transduction pathway of a cell of which a vertebrate
CC homologue of UNC-53 is a component comprising: (i) contacting an extract
CC of a cell with an antibody to the UNC-53 homologue; (ii) identifying an
CC antibody/homologue complex; and (iii) analysing such a complex to
CC identify any non-antibody protein bound to the complex. UNC-53 is a
CC signal transducing or signal integrator protein involved in controlling
CC directionality of cell migration and cell shape in *C. elegans*. Vertebrate
CC homologues of UNC-53 can be used to promote neuronal regeneration.
CC revascularisation or wound healing, to treat chronic neurodegenerative
CC diseases or acute traumatic injuries or fibrotic diseases. The present
CC sequence is used in the exemplification of the present invention.
SQ Sequence 1600 BP; 397 A; 409 C; 398 G; 396 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 2; Length 1600;
Best Local Similarity 100.0%; Pred. No. 1.2e-27;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTGCTCCCTGCTGTGTGGAGTCTGCTGAGTAGTGGCGAGCAAAATTTAAAGCTACA 60
DB 81 CTGCTCCCTGCTGTGTGGAGTCTGCTGAGTAGTGGCGAGCAAAATTTAAAGCTACA 140
QY 61 ACAAGGCAAGGCTTGACCGACAATTGCGATGAAGAAATCTGC 100
DB 141 ACAAGGCAAGGCTTGACCGACAATTGCGATGAAGAAATCTGC 180

RESULT 6
ADM41037
ID ADM41037 standard; DNA; 1782 BP.
XX
AC ADM41037;
XX
DT 17-JUN-2004 (first entry)
XX
DE Cytomegalovirus nucleotide sequence SEQ ID NO:5.
XX
KW engrafting foreign replacement cell; implanting foreign replacement cell;
KW growth; differentiation; drug development; vaccine development;
KW tissue transplantation; human disease study; cytomegalovirus; gene; ds.
XX
OS Cytomegalovirus.
XX
FN WO2004027029-A2.
XX
PD 01-APR-2004.
XX
PF 17-SEP-2003; 2003WO-US029251.
XX
PR 19-SEP-2002; 2002US-0411790P.
XX
PA (XIME-) XIMEREX INC.
XX
PI Beschornor WE, Sosa CE, Thompson SC;
XX
DR WPI; 2004-295402/27.
XX
PT Engrafting foreign replacement cells within a fetal non-human mammal,
PT useful in producing chimeric mammals, comprises selectively destroying
PT native cells in a tissue of a fetal non-human mammal host.
XX
PS Disclosure; SEQ ID NO 5; 48pp; English.
XX
CC The present invention describes a method for engrafting foreign
CC replacement cells within a fetal non-human mammal, which comprises
CC selectively destroying native cells in a tissue of a fetal non-human
CC mammal host, where the number of maternal cells of the same tissue is not
CC substantially reduced, and implanting foreign replacement cells in the
CC tissue of the fetal non-human mammal host, where the foreign replacement
CC cells replace destroyed cells of the tissue. The method is useful for
CC facilitating growth and differentiation of foreign cells within a
CC mammalian host, and for producing chimeric mammals that can be used to
CC develop new drugs and vaccine, factors, drugs and tissues for

CC transplantation, also useful to study human diseases. The present
CC sequence represents a nucleotide sequence given in the Sequence Listing
CC of the present invention but not mentioned further within the
CC specification.
SQ Sequence 1782 BP; 458 A; 399 C; 451 G; 474 T; 0 U; 0 Other;
Query Match 100.0%; Score 100; DB 12; Length 1782;
Best Local Similarity 100.0%; Pred. No. 1.2e-27;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTGCTCCCTGCTGTGTGGAGTCTGCTGAGTAGTGGCGAGCAAAATTTAAAGCTACA 60
DB 81 CTGCTCCCTGCTGTGTGGAGTCTGCTGAGTAGTGGCGAGCAAAATTTAAAGCTACA 140
QY 61 ACAAGGCAAGGCTTGACCGACAATTGCGATGAAGAAATCTGC 100
DB 141 ACAAGGCAAGGCTTGACCGACAATTGCGATGAAGAAATCTGC 180
RESULT 7
ADM41034
ID ADM41034 standard; DNA; 2241 BP.
XX
AC ADM41034;
XX
DT 17-JUN-2004 (first entry)
XX
DE Human nucleotide sequence SEQ ID NO:2.
XX
KW engrafting foreign replacement cell; implanting foreign replacement cell;
KW growth; differentiation; drug development; vaccine development;
KW tissue transplantation; human disease study; human; gene; ds.
XX
OS Homo sapiens.
XX
FN WO2004027029-A2.
XX
PD 01-APR-2004.
XX
PF 17-SEP-2003; 2003WO-US029251.
XX
PR 19-SEP-2002; 2002US-0411790P.
XX
PA (XIME-) XIMEREX INC.
XX
PI Beschornor WE, Sosa CE, Thompson SC;
XX
DR WPI; 2004-295402/27.
XX
PT Engrafting foreign replacement cells within a fetal non-human mammal,
PT useful in producing chimeric mammals, comprises selectively destroying
PT native cells in a tissue of a fetal non-human mammal host.
XX
PS Disclosure; SEQ ID NO 2; 48pp; English.
XX
CC The present invention describes a method for engrafting foreign
CC replacement cells within a fetal non-human mammal, which comprises
CC selectively destroying native cells in a tissue of a fetal non-human
CC mammal host, where the number of maternal cells of the same tissue is not
CC substantially reduced, and implanting foreign replacement cells in the
CC tissue of the fetal non-human mammal host, where the foreign replacement
CC cells replace destroyed cells of the tissue. The method is useful for
CC facilitating growth and differentiation of foreign cells within a
CC mammalian host, and for producing chimeric mammals that can be used to
CC develop new drugs and vaccine, factors, drugs and tissues for
CC transplantation, also useful to study human diseases. The present
CC sequence represents a nucleotide sequence given in the Sequence Listing
CC of the present invention but not mentioned further within the
CC specification.
SQ Sequence 2241 BP; 498 A; 630 C; 609 G; 504 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 12; Length 2241;
Best Local Similarity 100.0%; Pred. No. 1.3e-27;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCTCCCTGCTGTGTGTGAGGTGCTGAGTAGTGGCGAGCAAAATTTAAAGCTACA 60
DB 81 CTGCTCCCTGCTGTGTGTGAGGTGCTGAGTAGTGGCGAGCAAAATTTAAAGCTACA 140

QY 61 ACAAGGCAAGGCTTGACCGACAATTTGCATGAAGAATCTGC 100
DB 141 ACAAGGCAAGGCTTGACCGACAATTTGCATGAAGAATCTGC 180

RESULT 8
ABZ23249
ID ABZ23249 standard; DNA; 2245 BP.
XX AC ABZ23249;
XX DT 24-MAR-2003 (first entry)
XX DE Lac repressor operated p21-expression cassette and RSV-LTR promoter.
XX KW Lac repressor; p21; RSV; LTR promoter; cell cycle inhibitor protein;
XX KW protein production; anchorage-independent producer cell line; ss.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT promoter 1..563
FT /*tag= a
FT /note= "RSV-LTR promoter"
FT intron 564..1051
FT /*tag= b
FT /note= "SV40 small t antigen intron"
FT misc_feature 1052..1907
FT /*tag= c
FT /note= "p21 coding sequence"
FT polyA_signal 1908..2245
FT /*tag= d
FT /note= "thymidine kinase polyA site"

WO200299100-A2.
XX PN
XX PD
XX PF 03-JUN-2002; 2002WO-EP006054.
XX PR 01-JUN-2001; 2001GB-00013318.
XX PA (LONZ) LONZA BIOLOGICS PLC.
XX PI Al-Rubeai M, Shuttleworth J;
XX WPI; 2003-148669/14.
XX PT Producing recombinant protein, particularly for maximizing or enhancing
XX PT e.g. therapeutic protein production, by co-expressing protein with
XX PT recombinant cell cycle inhibitor protein (p21) in producer cell line.
XX PS Example 1; Page 15-16; 33pp; English.
XX CC The present sequence represents a lac repressor operated p21-expression
XX CC cassette comprising the Rous sarcoma virus (RSV)-LTR promoter. p21 is a
XX CC cell cycle inhibitor protein. The present sequence is used to produce
XX CC vectors for use in the method of the invention. The specification
XX CC describes a method for producing a protein, preferably a recombinant
XX CC protein, in a mammalian anchorage-independent producer cell line. The
XX CC method comprises co-expressing with the protein in the producer cell line
XX CC a recombinant cell cycle inhibitor protein (preferably p21). The method
XX CC is useful for producing a recombinant protein in a producer cell line.
XX CC This is particularly useful for maximizing or enhancing the production of
XX CC e.g. therapeutic proteins at an industrial scale

XX SQ Sequence 2245 BP; 532 A; 555 C; 625 G; 533 T; 0 U; 0 Other;
Query Match 100.0%; Score 100; DB 8; Length 2245;
Best Local Similarity 100.0%; Pred. No. 1.3e-27;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCTCCCTGCTGTGTGTGAGGTGCTGAGTAGTGGCGAGCAAAATTTAAAGCTACA 60
DB 46 CTGCTCCCTGCTGTGTGTGAGGTGCTGAGTAGTGGCGAGCAAAATTTAAAGCTACA 105

QY 61 ACAAGGCAAGGCTTGACCGACAATTTGCATGAAGAATCTGC 100
DB 106 ACAAGGCAAGGCTTGACCGACAATTTGCATGAAGAATCTGC 145

RESULT 9
ADM41036
ID ADM41036 standard; DNA; 2294 BP.
XX AC ADM41036;
XX DT 17-JUN-2004 (first entry)
XX DE Cytomegalovirus nucleotide sequence SEQ ID NO:4.
XX KW engrafting foreign replacement cell; implanting foreign replacement cell;
XX KW growth; differentiation; drug development; vaccine development;
XX KW tissue transplantation; human disease study; cytomegalovirus; gene; ds.
XX OS Cytomegalovirus.
XX PN WO2004027029-A2.
XX PD 01-APR-2004.
XX PF 17-SEP-2003; 2003WO-US029251.
XX PR 19-SEP-2002; 2002US-0411790P.
XX PA (XIME-) XIMEREX INC.
XX PI Beschoner WE, Sosa CE, Thompson SC;
XX WPI; 2004-295402/27.
XX PT Engrafting foreign replacement cells within a fetal non-human mammal,
XX PT useful in producing chimeric mammals, comprises selectively destroying
XX PS native cells in a tissue of a fetal non-human mammal host.
XX PS Disclosure; SEQ ID NO 4; 48pp; English.
XX CC The present invention describes a method for engrafting foreign
XX CC replacement cells within a foetal non-human mammal, which comprises
XX CC selectively destroying native cells in a tissue of a foetal non-human
XX CC mammal host, where the number of maternal cells of the same tissue is not
XX CC substantially reduced, and implanting foreign replacement cells in the
XX CC tissue of the fetal non-human mammal host, where the foreign replacement
XX CC cells replace destroyed cells of the tissue. The method is useful for
XX CC facilitating growth and differentiation of foreign cells within a
XX CC mammalian host, and for producing chimeric mammals that can be used to
XX CC develop new drugs and vaccine, factors, drugs and tissues for
XX CC transplantation, also useful to study human diseases. The present
XX CC sequence represents a nucleotide sequence given in the Sequence Listing
XX CC of the present invention but not mentioned further within the
XX CC specification.

XX SQ Sequence 2294 BP; 466 A; 680 C; 641 G; 507 T; 0 U; 0 Other;
Query Match 100.0%; Score 100; DB 12; Length 2294;
Best Local Similarity 100.0%; Pred. No. 1.3e-27;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTTGTGTGAGGTGCTGAGTAGTGCGGAGCAAAATTTAAGCTACA 60
 |||||
 Db 81 CTGCTCCCTGCTTGTGTGAGGTGCTGAGTAGTGCGGAGCAAAATTTAAGCTACA 140
 |||||

Qy 61 ACAAGGCAAGGCTTGACCGACAATTCGATGAAGAATCTGC 100
 |||||
 Db 141 ACAAGGCAAGGCTTGACCGACAATTCGATGAAGAATCTGC 180
 |||||

RESULT 10
 AAD02037
 ID AAD02037 standard; DNA; 2426 BP.
 XX
 AC AAD02037;
 XX
 DT 11-SEP-2003 (revised)
 DT 26-MAR-2001 (first entry)
 XX
 DE Plasmid pNG3/RC/CPG2-Thy1 comprising CPG2 DNA with rat thy1 gene.
 XX
 KW Carboxypeptidase G2; CPG2; gene directed enzyme produg therapy; GDEPT;
 KW glycosylphosphatidylinositol; GPI; cancer; therapy; rat; Thy1 gene;
 KW plasmid; ds.
 XX
 OS Rattus sp.
 OS Bacteria.
 OS Chimeric.
 XX
 PN WO200066752-A2.
 XX
 PD 09-NOV-2000.
 XX
 PF 28-APR-2000; 2000WO-GB001640.
 XX
 PR 01-MAY-1999; 99GB-00010077.
 XX
 PA (ASTR) ASTRAZENECA AB.
 PA (UYMA-) UNIV VICTORIA MANCHESTER.
 XX
 PI Castro MG, Emery SC, Lowenstein PR;
 PI WPI; 2001-015983/02.
 XX
 DR Gene directed enzyme produg therapy using post translational
 PT glycosylphosphatidylinositol addition to produg activating enzyme to
 PT enable anchorage of enzyme at cell surface for cancer therapy.
 XX
 PS Example 1e; Page 59-60; 60pp; English.
 XX
 CC The present invention relates to a gene directed enzyme produg therapy
 CC (GDEPT) using post translational glycosylphosphatidylinositol (GPI)
 CC addition to a produg activating enzyme which enables anchorage of the
 CC enzyme at the cell surface. Carboxypeptidase G2 (CPG2) is a preferred
 CC produg activating enzyme. The invention also relates to an expression
 CC vector for expression of a GPI enzyme hybrid capable of anchorage to the
 CC surface of a mammalian cell. The expression vector comprise
 CC polynucleotide sequences encoding a signal peptide, an enzyme capable of
 CC activating a produg, and a post-translational GPI addition motif. The
 CC expression vector is useful in the manufacture of a medicament for cancer
 CC therapy in a mammalian host. The present DNA sequence is a plasmid
 CC pNG3/RC/CPG2-Thy1 comprising CPG2 nucleic acid sequence with the last
 CC standardise OS field)
 XX
 SQ Sequence 2426 BP; 557 A; 705 C; 668 G; 495 T; 0 U; 1 Other;

Query Match 100.0%; Score 100; DB 4; Length 2426;
 Best Local Similarity 100.0%; Pred. No. 1.4e-27;
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 61 ACAAGGCAAGGCTTGACCGACAATTCGATGAAGAATCTGC 100
 |||||
 Db 129 ACAAGGCAAGGCTTGACCGACAATTCGATGAAGAATCTGC 168
 |||||

RESULT 11
 AAD02036
 ID AAD02036 standard; DNA; 2427 BP.
 XX
 AC AAD02036;
 XX
 DT 11-SEP-2003 (revised)
 DT 26-MAR-2001 (first entry)
 XX
 DE Plasmid pNG3/RC/CPG2 (Q3) -Thy1 comprising CPG2 variant with rat thy1 gene.
 XX
 KW Carboxypeptidase G2; CPG2; gene directed enzyme produg therapy; GDEPT;
 KW glycosylphosphatidylinositol; GPI; cancer; therapy; rat; Thy1 gene;
 KW CPG2 (Q3) variant; plasmid; ds.
 XX
 OS Rattus sp.
 OS Bacteria.
 OS Chimeric.
 XX
 PN WO200066752-A2.
 XX
 PD 09-NOV-2000.
 XX
 PF 28-APR-2000; 2000WO-GB001640.
 XX
 PR 01-MAY-1999; 99GB-00010077.
 XX
 PA (ASTR) ASTRAZENECA AB.
 PA (UYMA-) UNIV VICTORIA MANCHESTER.
 XX
 PI Castro MG, Emery SC, Lowenstein PR;
 PI WPI; 2001-015983/02.
 XX
 DR Gene directed enzyme produg therapy using post translational
 PT glycosylphosphatidylinositol addition to produg activating enzyme to
 PT enable anchorage of enzyme at cell surface for cancer therapy.
 XX
 PS Example 1e; Page 59; 60pp; English.
 XX
 CC The present invention relates to a gene directed enzyme produg therapy
 CC (GDEPT) using post translational glycosylphosphatidylinositol (GPI)
 CC addition to a produg activating enzyme which enables anchorage of the
 CC enzyme at the cell surface. Carboxypeptidase G2 (CPG2) is a preferred
 CC produg activating enzyme. The invention also relates to an expression
 CC vector for expression of a GPI enzyme hybrid capable of anchorage to the
 CC surface of a mammalian cell. The expression vector comprise
 CC polynucleotide sequences encoding a signal peptide, an enzyme capable of
 CC activating a produg, and a post-translational GPI addition motif. The
 CC expression vector is useful in the manufacture of a medicament for cancer
 CC therapy in a mammalian host. The present DNA sequence is a plasmid
 CC pNG3/RC/CPG2 (Q3) comprising CPG2 variant CPG2 (Q3) and the last exon of
 CC rat Thy-1 at the 3' end. (Updated on 11-SEP-2003 to standardise OS field)
 XX
 SQ Sequence 2427 BP; 555 A; 706 C; 670 G; 495 T; 0 U; 1 Other;

Query Match 100.0%; Score 100; DB 4; Length 2427;
 Best Local Similarity 100.0%; Pred. No. 1.4e-27;
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCCTGCTTGTGTGAGGTGCTGAGTAGTGCGGAGCAAAATTTAAGCTACA 60
 |||||
 Db 70 CTGCTCCCTGCTTGTGTGAGGTGCTGAGTAGTGCGGAGCAAAATTTAAGCTACA 129
 |||||

Qy 61 ACAAGGCAAGGCTTGACCGACAATTCGATGAAGAATCTGC 100
 |||||
 Db 130 ACAAGGCAAGGCTTGACCGACAATTCGATGAAGAATCTGC 169
 |||||

RESULT 12
 AAT62937
 ID AAT62937 standard; DNA; 3400 BP.
 XX AC AAT62937;
 XX DT 17-OCT-2003 (revised)
 XX DT 16-JUN-1997 (first entry)
 XX DE 3F4 human G2/G4 chimeric antibody expression plasmid insert.
 XX KW Xenotransplantation; graft rejection; cell interaction; pig;
 KW vascular cell adhesion molecule; VCAM; monoclonal antibody;
 XX chimeric antibody; diagnosis; ss.
 OS Homo; sapiens.
 OS Mus sp.
 OS Chimeric.
 XX FH Key
 FT exon Location/Qualifiers
 FT 903..1055
 FT /tag= a
 FT intron 1056..1285
 FT /tag= b
 FT exon 1286..2055
 FT /tag= c
 FT /codon_start= 1350
 FT intron 2056..2447
 FT /tag= d
 FT exon 2448..2483
 FT /tag= e
 FT intron 2484..2601
 FT /tag= f
 FT exon 2602..2928
 FT /tag= g
 FT intron 2929..3025
 FT /tag= h
 FT exon 3026..3348
 FT /tag= i
 XX WO9711971-Al.
 PN 03-APR-1997.
 PD 27-SEP-1996; 96WO-US015575.
 XX 28-SEP-1995; 95US-0004489P.
 PR 26-SEP-1996; 96US-00004489.
 XX (ALEX-) ALEXION PHARM INC.
 PA Mueller JP, Evans MJ, Mueller EE, Rollins S, Rother RP, Matis LA;
 PI WPI; 1997-212855/19.
 XX P-PSDB; AAW14940.
 DR Antibodies binding to porcine but not human cell interaction proteins -
 PT useful to treat and assay for rejection of xenografted porcine organs,
 PT tissues or cells.
 XX Disclosure; Page 58-61; 105pp; English.
 PS A DNA sequence (AAT62937) comprises a 3F4 human G2/G4 (see also AAT62936)
 CC chimeric antibody expression plasmid insert sequence. The chimeric
 CC antibody (AAW14940) is specific for porcine vascular cell adhesion
 CC molecule (VCAM) and is useful for diagnosing human rejection of porcine
 CC xenotransplants and for improving xenotransplantation of porcine cells,
 CC tissues and organs into human recipients. (Updated on 17-OCT-2003 to
 CC standardise OS field)
 XX Sequence 3400 BP; 759 A; 1012 C; 909 G; 720 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 2; Length 3400;
 Best Local Similarity 100.0%; Pred. No. 1.5e-27;
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CTGCTCCCTGCTTGTGTGGAGTGCCTGAGTAGTGGCGAGCAAAATTTAAGCTACA 60
 |||||||
 Db 148 CTGCTCCCTGCTTGTGTGGAGTGCCTGAGTAGTGGCGAGCAAAATTTAAGCTACA 207
 |||||||
 QY 61 ACAAGGCAAGCTTGCACCGACAATTGCATGAAGAATCTGC 100
 |||||||
 Db 208 ACAAGGCAAGCTTGCACCGACAATTGCATGAAGAATCTGC 247
 |||||||
 RESULT 13
 AAT62932
 ID AAT62932 standard; DNA; 3400 BP.
 XX AC AAT62932;
 XX DT 17-OCT-2003 (revised)
 XX DT 16-JUN-1997 (first entry)
 XX DE 2A2 human G2/G4 chimeric antibody expression plasmid insert.
 XX KW Xenotransplantation; graft rejection; cell interaction; pig;
 KW vascular cell adhesion molecule; VCAM; monoclonal antibody;
 XX chimeric antibody; diagnosis; ss.
 OS Homo; sapiens.
 OS Mus sp.
 OS Chimeric.
 XX FH Key
 FT exon Location/Qualifiers
 FT 903..1055
 FT /tag= a
 FT intron 1056..1285
 FT /tag= b
 FT exon 1286..2020
 FT /tag= c
 FT /codon_start= 1318
 FT intron 2021..2412
 FT /tag= d
 FT exon 2413..2448
 FT /tag= e
 FT intron 2449..2566
 FT /tag= f
 FT exon 2567..2983
 FT /tag= g
 FT intron 2984..2990
 FT /tag= h
 FT exon 2991..3313
 FT /tag= i
 XX WO9711971-Al.
 PN 03-APR-1997.
 PD 27-SEP-1996; 96WO-US015575.
 XX 28-SEP-1995; 95US-0004489P.
 PR 26-SEP-1996; 96US-00004489.
 XX (ALEX-) ALEXION PHARM INC.
 PA Mueller JP, Evans MJ, Mueller EE, Rollins S, Rother RP, Matis LA;
 PI WPI; 1997-212855/19.
 XX P-PSDB; AAW14934.
 DR Antibodies binding to porcine but not human cell interaction proteins -
 PT useful to treat and assay for rejection of xenografted porcine organs,
 PT tissues or cells.

```
XX PS Disclosure; Page 44-47; 105pp; English.
XX CC A DNA sequence (AAT62932) comprises a 2A2 human G2/G4 (see also AAT62931)
XX CC chimeric antibody expression plasmid insert sequence. The chimeric
XX CC antibody (AAW14934) is specific for porcine vascular cell adhesion
XX CC molecule (VCAM) and is useful for diagnosing human rejection of porcine
XX CC xenotransplants and for improving xenotransplantation of porcine cells,
XX CC tissues and organs into human recipients. (Updated on 17-OCT-2003 to
XX CC standardise OS field)
XX SQ Sequence 3400 BP; 776 A; 993 C; 899 G; 732 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 100; DB 2; Length 3400;
XX Best Local Similarity 100.0%; Pred. No. 1.5e-27;
XX Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 CTGCTCCCTGCTGTGTGGAGTGCCTGAGTAGTGCAGCAGCAAAATTTAAGCTACA 60
XX Db CTGCTCCCTGCTGTGTGGAGTGCCTGAGTAGTGCAGCAGCAAAATTTAAGCTACA 207
XX
XX QY 61 ACAAGGCAAGCTTGACCGACAATTGCATGAAGAATCTGC 100
XX Db 208 ACAAGGCAAGCTTGACCGACAATTGCATGAAGAATCTGC 247
XX
XX RESULT 14
XX AAV40006
XX ID AAV40006 standard; DNA; 3853 BP.
XX AC AAV40006;
XX XX
XX DT 27-AUG-2003 (revised)
XX DT 15-FEB-1999 (first entry)
XX XX
XX DE Plasmid pCTM.
XX XX
XX KW E2F; transcription factor; human; retinoblastoma protein RB;
XX KW bladder cancer; retinosis; angioplasty; diabetic retinopathy;
XX KW thyroid hyperplasia; Grave's disease; psoriasis;
XX KW benign prostatic hypertrophy; Li-Fraumeni syndrome;
XX KW peripheral vascular disease; therapy; plasmid pCTM; ss.
XX XX
XX OS Human cytomegalovirus.
XX OS macadenovirus.
XX OS unidentified bacteriophage; T7.
XX OS unidentified bacteriophage; SP6.
XX OS Macaca mulatta; polyoma virus.
XX OS Bos taurus.
XX OS Chimeric.
XX XX
XX FH Key Location/Qualifiers
XX FT promoter 209..864
XX FT /tag= a
XX FT /note= "CMV promoter"
XX FT misc_feature 907..1131
XX FT /tag= b
XX FT /function= "tripartite leader sequence"
XX FT promoter 1132..1149
XX FT /tag= c
XX FT /note= "SP6 promoter"
XX FT misc_feature 1679..3853
XX FT /tag= d
XX FT /note= "pUC19 backbone H3 to AatII"
XX FT CDS complement (2857..3717)
XX FT /tag= e
XX FT /note= "AMP-ORF"
XX XX
XX XX WO9821228-A1.
XX XX 22-MAY-1998.
XX XX
XX PF 13-NOV-1997; 97WO-US021821.
```

```
XX PS 15-NOV-1996; 96US-00751517.
XX PR 14-FEB-1997; 97US-00801092.
XX XX
XX PA (CANJ-) CANJI INC.
XX XX
XX PI Antelman D, Gregory RJ, Wills KN;
XX XX
XX DR WPI; 1998-297858/26.
XX XX
XX PT New fusion polypeptide of, e.g. transcription factor - used to treat,
XX PT e.g. hyper-proliferative disease such as cancer and restenosis.
XX XX
XX PS Example 1; Fig 4; 91pp; English.
XX XX
XX CC This is the nucleotide sequence of pCTM, a plasmid which contains a CMV
XX CC promoter, a tripartite adenovirus leader flanked by T7 and SP6 promoters,
XX CC and a multiple cloning site with a bovine growth hormone polyA site and
XX CC downstream SV40 polyA site. It has been used as a vector for the
XX CC expression of fusion proteins of the invention that comprise
XX CC retinoblastoma protein (BP, see AAW62465) and E2F transcription factor
XX CC (see AAW62464). Such fusion proteins, particularly expressed from gene
XX CC therapy vectors, are used to treat hyperproliferative conditions,
XX CC specifically cancer (particularly of the bladder) or restenosis. They are
XX CC more effective in representing transcription of the E2F promoter than RA
XX CC alone and cause cell-cycle arrest in a variety of cells. (Updated on 27-
XX CC AUG-2003 to correct OS field.)
XX XX
XX SQ Sequence 3853 BP; 936 A; 986 C; 942 G; 989 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 100; DB 2; Length 3853;
XX Best Local Similarity 100.0%; Pred. No. 1.6e-27;
XX Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 CTGCTCCCTGCTGTGTGGAGTGCCTGAGTAGTGCAGCAGCAAAATTTAAGCTACA 60
XX Db CTGCTCCCTGCTGTGTGGAGTGCCTGAGTAGTGCAGCAGCAAAATTTAAGCTACA 140
XX
XX QY 61 ACAAGGCAAGCTTGACCGACAATTGCATGAAGAATCTGC 100
XX Db 141 ACAAGGCAAGCTTGACCGACAATTGCATGAAGAATCTGC 180
XX
XX RESULT 15
XX AAT90695
XX ID AAT90695 standard; DNA; 3925 BP.
XX AC AAT90695;
XX XX
XX DT 05-JAN-1998 (first entry)
XX XX
XX XX Plasmid CMV10A1 coding sequence.
XX XX
XX KW Packaging-deficient construct; viral gag-pol gene; packaging cell line;
XX KW moloney murine leukaemia virus; MoMLV; viral env gene; helper construct;
XX KW gene therapy; humna cytomegalovirus; promoter; ss.
XX OS Synthetic.
XX XX
XX XX WO9708330-A1.
XX XX
XX XX 06-MAR-1997.
XX XX
XX XX 23-AUG-1996; 96WO-GB002061.
XX XX
XX XX 23-AUG-1995; 95GB-00017263.
XX XX
XX XX (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
XX XX
XX XX Collins MKL, Weiss RA, Takeuchi Y, Cosset F;
XX XX WPI; 1997-179287/16.
XX XX
```

PT Selectable retroviral packaging cell lines and expression constructs -
PT comprise selectable gene downstream of gene of interest, are selectable
due to the in-efficiency associated with translation re-initiation.

PS Claim 23; Fig 13; 79pp; English.

XX
XX
CC This sequence represents the recombinant expression plasmid CMV10A. This
CC sequence is a packaging-deficient construct having a viral env gene (in
CC this case from moloney murine leukaemia virus under hCMV promoter
CC control) and a selectable marker (SM). It is an example of a recombinant
CC expression vector (REV) of the invention, used to create a packaging cell
CC line. The REVs of the invention comprise a gene of interest (GOI) and a
CC SM gene. The SM gene is arranged downstream of the GOI and a GOI
CC associated stop codon is spaced from a start codon of the SM gene to
CC ensure that the SM protein is expressed as a result of translation
CC reinitiation. The cell lines are transformed with two REVs, both are
CC replication deficient, one contains the viral gag-pol gene, the other the
CC viral env gene. By using helper constructs, such as the REV's, which are
CC directly selectable and which provide for high expression of the viral
CC gene, high titre retroviral vectors may be obtained. The packaging cell
CC lines are useful for gene therapy. Prior packaging cell lines using full
CC length retroviral genomes as helper genomes were isolated by
CC cotransfecting them with plasmids encoding selectable markers. However,
CC the helper functions can be lost during the passages of the cells in
CC culture and the current packaging systems provide limited titres of
CC infectious retroviral vectors. Co-transfection with a plasmid encoding a
CC SM does not directly select the best gag-pol-env-expressing cells. The
CC new retroviral packaging cell lines overcome these problems

XX SQ Sequence 3925 BP; 963 A; 1001 C; 959 G; 998 T; 0 U; 4 Other;

Query Match 100.0%; Score 100; DB 2; Length 3925;
Best Local Similarity 100.0%; Pred. No. 1.6e-27;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CTGCTCCCTGCTTGTGTGAGTCTGCTGAGTAGTCGGCGAGCAAAATTTAAGCTACA 60
Db 70 CTGCTCCCTGCTTGTGTGAGTCTGCTGAGTAGTCGGCGAGCAAAATTTAAGCTACA 129
Qy 61 ACAAGGCAAGGCTTGACCGACAATTGCATGAGATCTGC 100
Db 130 ACAAGGCAAGGCTTGACCGACAATTGCATGAGATCTGC 169

Search completed: July 14, 2005, 07:01:48
Job time : 141.038 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 05:15:57 ; Search time 952.146 Seconds
(without alignments)
3997.736 Million cell updates/sec

Title: US-09-482-682-65_COPY_1_100
Perfect score: 100
Sequence: 1 ctgtccctctgtgtgtt.....caattgcatgaagaattctgc 100

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*

1: gb_est1.*
2: gb_est2.*
3: gb_hc.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gss1.*
9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100	100.0	602	8	B67169 CPG0047A Cp
2	10.4	30.4	829	4	BI333630 602997459
3	30.2	30.2	823	6	CD655614 AGENCOURT
4	29.8	29.8	1165	8	CC242469 CH261-11P
5	29.4	29.4	401	6	CB387202 OSTF076E6
6	29.4	29.4	531	5	BQ310441 MR0-BT450
7	29.4	29.4	754	8	AQ946479 Sheared D
8	29.2	29.2	426	9	CC888514 SALK 1519
9	29	29.0	657	7	CK086063 RG11-C07
10	29	29.0	877	9	AL225351 Tetradon
11	28.6	28.6	340	7	F32722 HSPD25699 H
12	28.6	28.6	408	1	AA962465 co91e05.s
13	28.6	28.6	436	7	CN386744 328755673
14	28.6	28.6	514	6	CB161201 K-EST0221
15	28.6	28.6	530	6	CB161182 K-EST0220
16	28.6	28.6	534	2	AW500392 UI-HF-BN0
17	28.6	28.6	550	1	AA984313 am83h04.s
18	28.6	28.6	555	7	CR537056 DKEP2p459D
19	28.6	28.6	583	7	CN386728 170005326
20	28.6	28.6	585	4	BG993413 MR3-HT099
21	28.6	28.6	625	4	BI113747 602860946
22	28.6	28.6	626	7	CN386724 170006000
23	28.6	28.6	641	2	AW955076 EST367146
24	28.6	28.6	651	7	CN386730 170005316

C 25	28.6	28.6	680	7	CN386731
C 26	28.6	28.6	682	7	CN386746
C 27	28.6	28.6	687	7	CN386717
C 28	28.6	28.6	710	2	BF309673
C 29	28.6	28.6	724	7	CN386780
C 30	28.6	28.6	770	6	CD654097
C 31	28.6	28.6	777	7	CN386742
C 32	28.6	28.6	801	6	CD656906
C 33	28.6	28.6	814	4	BG820298
C 34	28.6	28.6	823	6	CB996419
C 35	28.6	28.6	827	4	BG387788
C 36	28.6	28.6	830	6	CD643822
C 37	28.6	28.6	842	5	BU170446
C 38	28.6	28.6	851	6	CB993607
C 39	28.6	28.6	852	5	BQ222104
C 40	28.6	28.6	856	6	CD656102
C 41	28.6	28.6	909	5	BU189860
C 42	28.6	28.6	1025	4	BM477450
C 43	28.4	28.4	579	6	CB239722
C 44	28.4	28.4	733	7	CF667137
C 45	28.4	28.4	1015	6	BY703355

ALIGNMENTS

RESULT 1
B67169
LOCUS B67169 602 bp DNA linear GSS 12-MAY-2000
DEFINITION CPG0047A CpiOWAgDNA2 Cryptosporidium parvum genomic, genomic survey sequence.
ACCESSION B67169
VERSION B67169.1 GI:2642750
KEYWORDS GSS.
SOURCE Cryptosporidium parvum
ORGANISM Cryptosporidium parvum
REFERENCE 1 (bases 1 to 602)
AUTHORS Strong, W.B. and Nelson, R.G.
TITLE Preliminary profile of the Cryptosporidium parvum genome: an expressed sequence tag and genome survey sequence analysis
JOURNAL Mol. Biochem. Parasitol. 107 (1), 1-32 (2000)
MEDLINE 20183851
PUBMED 10717299
COMMENT Contact: Nelson, R. G.
Depts. of Medicine & Pharmaceutical Chemistry
San Francisco General Hospital-University of California, San Francisco
Box 0811, San Francisco, CA 94143-0811, USA
Tel: 415 206 8846
Fax: 415 206 3353
Email: malariaditsea.ucsf.edu
Submitted sequence has been edited to remove vector sequences 5' to the insert, to correct miscalled bases and assign uncalled (N) bases throughout the sequence, and to terminate when base-calling became ambiguous.
Seq primer: T7
Class: Shotgun
High quality sequence stop: 602.
Location/Qualifiers
1. .602
/organism="Cryptosporidium parvum"
/mol_type="genomic DNA"
/strain="IOWA"
/db_xref="taxon:5807"
/lab_host="E. coli XL2 Blue MRF"
/clone_lib="CpiOWAgDNA2"
/note="Vector: PCR-Script Amp SK+; Site 1: SrfI; C. parvum (IOWA isolate) genomic DNA was hydrodynamically sheared to produce fragments having a tight size distribution between 2-4 kb by Dr. Yvonne Thorstenson of the Stanford DNA Sequencing and Technology Center

(http://sequence-www.stanford.edu/group/techdev/shear.htm)

The randomly sheared gDNA was chromatographed on Sephacryl S-400 to remove any small fragments and DNA eluting in the void volume was blunt-ended with T4 DNA Polymerase, treated with alkaline phosphatase to prevent the ligation of multiple fragments, ligated to Srf I-digested PCR-Script Amp (SK+) vector and transformed into E. coli strain XL10 Gold. Recombinant clones from the first plating of the library were selected for sequence analysis using T3 and T7 primers."

ORIGIN

Query Match 100.0%; Score 100; DB 8; Length 602;
 Best Local Similarity 100.0%; Pred. No. 8e-24;
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCTCCTGCTGTTGTTGGAGTCTGCTGAGTAGTGGCGAGCAAAATTTAAGCTACA 60
 |||||
 Db 41 CTGCTCCTGCTGTTGTTGGAGTCTGCTGAGTAGTGGCGAGCAAAATTTAAGCTACA 100
 |||||

Qy 61 ACAAGCAAGCTTGACGCAATTTGATGAAGAAATCTGC 100
 |||||
 Db 101 ACAAGCAAGCTTGACGCAATTTGATGAAGAAATCTGC 140
 |||||

RESULT 2

B1333630/c
 LOCUS 602997459F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5139470 5',
 DEFINITION mRNA sequence.
 ACCESSION B1333630
 VERSION B1333630.1 GI:15018287
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 829)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: Incyte Genomics, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1AM11343 row: f column: 15
 High quality sequence stop: 788.
 Location/Qualifiers
 1..829
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5139470"
 /tissue_type="cervical carcinoma cell line"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_12"
 /note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr. Average insert size 1.4 kb. Library prepared by Life Technologies."

REFERENCE

AUTHORS
 TITLE
 JOURNAL
 COMMENT

FEATURES

source
 1..829
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5139470"
 /tissue_type="cervical carcinoma cell line"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_12"
 /note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr. Average insert size 1.4 kb. Library prepared by Life Technologies."

ORIGIN

Query Match 30.4%; Score 30.4; DB 4; Length 829;
 Best Local Similarity 67.2%; Pred. No. 15;
 Matches 43; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy 19 TTGAGGTGCGTGTAGTGGCGAGCAAAATTTAAGCTACAACAGCGAGGCTTGACC 78
 |||||

Db 217 TTGGGCGTCCAGAAATTGTTGGTGAGCAAACTTCAAGTTGCTGCTGGGAAGTCTTGACT 158
 Qy 79 GACA 82
 |||||
 Db 157 GACA 154

RESULT 3

CD655614/c
 LOCUS 823 bp mRNA linear EST 18-JUN-2003
 DEFINITION (long) Homo sapiens cDNA clone IMAGE:30424285 5', mRNA sequence.
 ACCESSION CD655614
 VERSION CD655614.1 GI:31896113
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 823)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Irene Ginis and Mahendra Rao, NIA
 cDNA Library Preparation: Yulan Piao and Minoru Ko
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: NDAM506 row: k column: 14
 High quality sequence stop: 640.
 Location/Qualifiers
 1..823
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30424285"
 /tissue_type="Embryonic Stem cells"
 /cell_line="WA01"
 /lab_host="NIA Human H1 Embryonic Stem Cell cDNA Library (Long)"
 /note="Vector: pCMV-Sport6; Site 1: NotI; Site 2: SalI; This is a long-transcript enriched cDNA library (Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]) from WA01 cell line. Undifferentiated human ES cell line WA01/H1 was obtained from WiCell Research Institute, Inc., Madison, WI, cultured according to their instructions, on MSF feeders. They formed round colonies with defined edges and were positive for alkaline phosphatase, SSEA-4, OCT3, OCT4, REX1, UTP, TERT, SOX2, CX43 and CX45. They are negative for GATA2, GATA4, PDX1, NCAM, MSX1, FLT3, SSEA-1, TUBB3, NES, GFAP, and BOMES. When confluent (18-10 days after plating), the ES cells from 4 x 6 cm dishes were treated with 1 mg/ml collagenase, type IV (Invitrogen/GIBCO) for 5-10 min and gently scraped off with 5 ml pipette. RNA was purified with TRIzol Reagent from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]) Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pGACTAGTCTAGATCGGAGCGCGCTTTTCTTTT-3'] from 3.4g of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker LL-Sal4, purified by phenol/chloroform extraction, and separated from free linkers by Centricon-100 column. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase

FEATURES

source
 1..823
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30424285"
 /tissue_type="Embryonic Stem cells"
 /cell_line="WA01"
 /lab_host="NIA Human H1 Embryonic Stem Cell cDNA Library (Long)"
 /note="Vector: pCMV-Sport6; Site 1: NotI; Site 2: SalI; This is a long-transcript enriched cDNA library (Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]) from WA01 cell line. Undifferentiated human ES cell line WA01/H1 was obtained from WiCell Research Institute, Inc., Madison, WI, cultured according to their instructions, on MSF feeders. They formed round colonies with defined edges and were positive for alkaline phosphatase, SSEA-4, OCT3, OCT4, REX1, UTP, TERT, SOX2, CX43 and CX45. They are negative for GATA2, GATA4, PDX1, NCAM, MSX1, FLT3, SSEA-1, TUBB3, NES, GFAP, and BOMES. When confluent (18-10 days after plating), the ES cells from 4 x 6 cm dishes were treated with 1 mg/ml collagenase, type IV (Invitrogen/GIBCO) for 5-10 min and gently scraped off with 5 ml pipette. RNA was purified with TRIzol Reagent from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]) Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pGACTAGTCTAGATCGGAGCGCGCTTTTCTTTT-3'] from 3.4g of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker LL-Sal4, purified by phenol/chloroform extraction, and separated from free linkers by Centricon-100 column. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase

Salk Institute Genomic Analysis Laboratory (SIGnAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: eker@salik.edu
This is single pass sequence recovered from the left border of
TDNA. This sequence lies within an annotated intron of At1g03960.
Class: TDNA tagged.

FEATURES

Location/Qualifiers

source

1. 426
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Col-0"
/db_xref="taxon:3702"
/clone="SALK151964.54.50.x"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna_protocols.html"

ORIGIN

Query Match 29.2%; Score 29.2; DB 9; Length 426;
Best Local Similarity 64.2%; Pred. No. 35;
Matches 43; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

Qy 30 TGAGTAGTCGCGAGCAAAATTTAGCTACACAGGCGCTTGACCGACAATTCGAT 89

Db 76 TTATTAGTTGGTGTTCAGANTTTAGCATCATCAAGACTTGACCTACAGAACAT 17

Qy 90 GAAGAAT 96

Db 16 GAAAAAT 10

RESULT 9

CK086063/c
LOCUS CK086063 657 bp mRNA linear EST 01-DEC-2003
DEFINITION RG11_C07 Cucumber leaf Cucumis sativus cDNA, mRNA sequence.

ACCESSION CK086063

VERSION CK086063.1 GI:38571123

KEYWORDS EST.

SOURCE Cucumis sativus (cucumber)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Cucurbitales; Cucurbitaceae; Cucumis.

REFERENCE

1 (bases 1 to 657)

Grumet, R. and McGrath, M.

Development of genomic tools for cucumber (*Cucumis sativus* L.)

Unpublished (2003)

CONTACT: Rebecca Grumet

Rebecca Grumet

Michigan State University

Horticulture Department, Michigan State University, East Lansing,

MI 48824, USA

Tel: 517 353 0890

Fax: 517 355 5191 x431

Email: grumet@msu.edu

Plate: RG11 row: C column: 07.

Location/Qualifiers

source

1. 657
/organism="Cucumis sativus"
/mol_type="mRNA"
/strain="Straight 8"
/db_xref="taxon:3659"
/sex="monoecious"
/clone_lib="Cucumber leaf"
/note="Vector: pAD-GAL4; Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Query Match 29.0%; Score 29; DB 7; Length 657;
Best Local Similarity 63.8%; Pred. No. 44;
Matches 44; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Qy 31 GAGTAGTCGCGAGCAAAATTTAGCTACACAGGCGCTTGACCGACAATTCGATG 90

Db 267 GAGGAGTGCAGGAAGCAAACTGAAGCCAGAAATGAAGAGAGGCTGGACTTCAAGGTCAAG 208

Qy 91 AAGAATCTG 99

Db 207 AAAAATTTG 199

RESULT 10

CNS032VI/c

LOCUS CNS032VI 877 bp DNA linear GSS 01-SEP-2000

DEFINITION Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone
207F02 of library G from Tetraodon nigroviridis, genomic survey
sequence.

ACCESSION AL225351

VERSION AL225351.1 GI:7884242

KEYWORDS GSS; genome survey sequence.

SOURCE Tetraodon nigroviridis

ORGANISM Tetraodon nigroviridis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Acanthopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontoidea; Tetraodontidae; Tetraodon.

REFERENCE

1

Roest Crolius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C.,

Barnot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, F.,

Saurin, W. and Weissenbach, J.

Estimate of human gene number provided by genome-wide analysis

using Tetraodon nigroviridis DNA sequence

Nat. Genet. 25 (2), 235-238 (2000)

20296633

10835645

REFERENCE 2

Roest Crolius, H., Jaillon, O., Dasilva, C., Ozouf-Costaz, C.,

Fizames, C., Fischer, C., Bouneau, L., Billault, A., Quetier, F.,

Saurin, W., Bernot, A. and Weissenbach, J.

Characterization and repeat analysis of the compact genome of the

freshwater pufferfish Tetraodon nigroviridis

Genome Res. 10 (7), 939-949 (2000)

20359837

10899143

REFERENCE 3 (bases 1 to 877)

Genoscope.

Direct Submission

Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :

BP 131 91006 EVRY cedex - FRANCE (E-mail : secref@genoscope.cns.fr)

- Web : www.genoscope.cns.fr

This sequence is a single read and was generated as part of a large

scale clone-end sequencing project of the Tetraodon nigroviridis

genome. For more information, please take a look at

<http://www.genoscope.cns.fr/Tetraodon>.

Location/Qualifiers

1. 877

/organism="Tetraodon nigroviridis"

/mol_type="genomic DNA"

/db_xref="taxon:99883"

/clone="207F02"

/clone_lib="G"

/notes="Genoscope sequence ID : COAG207DC01SP1-end ;

PUC-Ori"

ORIGIN

Query Match 29.0%; Score 29; DB 9; Length 877;
Best Local Similarity 63.8%; Pred. No. 47;
Matches 44; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Qy 14 GTGTGTGGAGGTCGCTGAGTAGTCGCGAGCAAAATTTAAGCTACACAGGCAAGGCT 73


```

TITLE      Transcriptome characterization elucidates signaling networks that
JOURNAL    control human ES cell growth and differentiation
COMMENT    Nat. Biotechnol. 22 (6), 707-716 (2004)
Contact: Brandenberger R
Regenerative Medicine
Genon Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@genon.com
Insert Length: 436 Std Error: 0.00.
FEATURES   Location/Qualifiers
source     1..436
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /tissue_type="embryonic stem cells, embryoid bodies
            derived from H1, H7 and H9 cells"
            /clone_lib="GRN_EB"
            /notes="oligo dt primed, full-length enriched cDNA library
            from embryoid body outgrowths derived from hES cell lines
            H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
            conditions."

ORIGIN
Query Match      28.6%; Score 28.6; DB 7; Length 436;
Best Local Similarity 61.3%; Pred. No. 57;
Matches 46; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy 22 GAGTCGCTGAGTAGTGGCGAGCAAAATTTAAAGCTACAAAGGCAAGGCTTGACCGAC 81
Db 133 GAGCTAACTGAATTTGATGGGAGCAGCATTTAAACATATTCCTAGTCAAGGACAGGATGGG 74
Qy 82 AATTGCATGAAGAAT 96
Db 73 AAGTAAGTGAAGAAT 59

RESULT 14
LOCUS      CB161201
DEFINITION K-EST0221011 L18POOL1n1 Homo sapiens cDNA clone L18POOL1n1-16-H03
5', mRNA sequence.
ACCESSION  CB161201
VERSION     CB161201.1 GI:28147327
KEYWORDS   EST.
SOURCE      Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 514)
AUTHORS     Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
TITLE       21C Frontier Korean EST Project 2001
JOURNAL     Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 16, row: H column: 03
            High quality sequence stop: 514.
            Location/Qualifiers
FEATURES     source
            1..514
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clones="L18POOL1n1-16-H03"
            /cell_line="SNU-354+Cho-CK+Choi-HLK-3"
            /lab_host="Top10F"

TITLE      Transcriptome characterization elucidates signaling networks that
JOURNAL    control human ES cell growth and differentiation
COMMENT    Nat. Biotechnol. 22 (6), 707-716 (2004)
Contact: Brandenberger R
Regenerative Medicine
Genon Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@genon.com
Insert Length: 436 Std Error: 0.00.
FEATURES   Location/Qualifiers
source     1..436
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clones="L18POOL1n1-16-H03"
            /cell_line="SNU-354+Cho-CK+Choi-HLK-3"
            /lab_host="Top10F"

ORIGIN
Query Match      28.6%; Score 28.6; DB 6; Length 514;
Best Local Similarity 61.3%; Pred. No. 59;
Matches 46; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy 22 GAGTCGCTGAGTAGTGGCGAGCAAAATTTAAAGCTACAAAGGCAAGGCTTGACCGAC 81
Db 261 GAGCTAACTGAATTTGATGGGAGCAGCATTTAAACATATTCCTAGTCAAGGACAGGATGGG 320
Qy 82 AATTGCATGAAGAAT 96
Db 321 AAGTAAGTGAAGAAT 335

RESULT 15
LOCUS      CB161182
DEFINITION K-EST0220988 L18POOL1n1 Homo sapiens cDNA clone L18POOL1n1-16-F04
5', mRNA sequence.
ACCESSION  CB161182
VERSION     CB161182.1 GI:28147308
KEYWORDS   EST.
SOURCE      Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 530)
AUTHORS     Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
TITLE       21C Frontier Korean EST Project 2001
JOURNAL     Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 16 row: F column: 04
            High quality sequence stop: 530.
            Location/Qualifiers
FEATURES     source
            1..530
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clones="L18POOL1n1-16-F04"
            /cell_line="SNU-354+Cho-CK+Choi-HLK-3"
            /lab_host="Top10F"
            /clone_lib="L18POOL1n1"
            /notes="Organ: Liver; Vector: pT7T3-Pac; Site 1: EcoRI;
            Site 2: NotI; The library was contributed by the Soares
            laboratory and it was constructed as described by Bonaldo,
            M.F., Lennon, G. and Soares, M.B. (1996), Genome Research
            6(9): 791-806. RNA was prepared from harvested cell
            culture."

ORIGIN
Query Match      28.6%; Score 28.6; DB 6; Length 530;
Best Local Similarity 61.3%; Pred. No. 59;
Matches 46; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy 22 GAGTCGCTGAGTAGTGGCGAGCAAAATTTAAAGCTACAAAGGCAAGGCTTGACCGAC 81
Db 261 GAGCTAACTGAATTTGATGGGAGCAGCATTTAAACATATTCCTAGTCAAGGACAGGATGGG 320

```

```

/clone_lib="L18POOL1n1"
/notes="Organ: Liver; Vector: pT7T3-Pac; Site 1: EcoRI;
Site 2: NotI; The library was contributed by the Soares
laboratory and it was constructed as described by Bonaldo,
M.F., Lennon, G. and Soares, M.B. (1996), Genome Research
6(9): 791-806. RNA was prepared from harvested cell
culture."

ORIGIN
Query Match      28.6%; Score 28.6; DB 6; Length 514;
Best Local Similarity 61.3%; Pred. No. 59;
Matches 46; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy 22 GAGTCGCTGAGTAGTGGCGAGCAAAATTTAAAGCTACAAAGGCAAGGCTTGACCGAC 81
Db 261 GAGCTAACTGAATTTGATGGGAGCAGCATTTAAACATATTCCTAGTCAAGGACAGGATGGG 320
Qy 82 AATTGCATGAAGAAT 96
Db 321 AAGTAAGTGAAGAAT 335

RESULT 15
LOCUS      CB161182
DEFINITION K-EST0220988 L18POOL1n1 Homo sapiens cDNA clone L18POOL1n1-16-F04
5', mRNA sequence.
ACCESSION  CB161182
VERSION     CB161182.1 GI:28147308
KEYWORDS   EST.
SOURCE      Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 530)
AUTHORS     Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
TITLE       21C Frontier Korean EST Project 2001
JOURNAL     Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 16 row: F column: 04
            High quality sequence stop: 530.
            Location/Qualifiers
FEATURES     source
            1..530
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clones="L18POOL1n1-16-F04"
            /cell_line="SNU-354+Cho-CK+Choi-HLK-3"
            /lab_host="Top10F"
            /clone_lib="L18POOL1n1"
            /notes="Organ: Liver; Vector: pT7T3-Pac; Site 1: EcoRI;
            Site 2: NotI; The library was contributed by the Soares
            laboratory and it was constructed as described by Bonaldo,
            M.F., Lennon, G. and Soares, M.B. (1996), Genome Research
            6(9): 791-806. RNA was prepared from harvested cell
            culture."

ORIGIN
Query Match      28.6%; Score 28.6; DB 6; Length 530;
Best Local Similarity 61.3%; Pred. No. 59;
Matches 46; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy 22 GAGTCGCTGAGTAGTGGCGAGCAAAATTTAAAGCTACAAAGGCAAGGCTTGACCGAC 81
Db 261 GAGCTAACTGAATTTGATGGGAGCAGCATTTAAACATATTCCTAGTCAAGGACAGGATGGG 320

```

QY 82 AATTGCATGAAGAT 96
|||
Db 321 AAGTAAGTGAAGAT 335
|||

Search completed: July 14, 2005, 23:23:20
Job time : 952.146 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:39:07 ; Search time 756.618 Seconds
(without alignments)
6468.225 Million cell updates/sec

Title: US-09-482-682-65_COPY_4192_4292
Perfect score: 101
Sequence: 1 tgccttgccgcgtgctggac.....ccatggctgctgcgtgtgt 101

Scoring table:
IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	609	6	AX937037
2	101	100.0	609	6	AX952091
3	101	100.0	1973	14	OHVSAGA
4	101	100.0	2649	14	AF410858
5	101	100.0	2655	14	AF410857
6	101	100.0	2655	14	AF410859
7	101	100.0	2655	14	AF410860
8	101	100.0	2655	14	AF410861
9	101	100.0	3320	14	OHVGB
10	101	100.0	3323	14	OHVGC
11	101	100.0	3323	14	OHVCGD
12	101	100.0	3323	14	OHVHEPBA
13	101	100.0	7207	12	AX468486
14	101	100.0	7515	6	AX663053
15	101	100.0	7808	10	MAR87
16	101	100.0	8484	6	B268253
17	101	100.0	10469	6	AX641836
18	99.4	98.4	2655	14	AF410855
19	99.4	98.4	2655	14	AF410856

20	96.2	95.2	3308	14	OHVSURCOR	M90520 Woodchuck h
21	94.6	93.7	592	6	AR136166	AR136166 Sequence
22	94.6	93.7	592	6	AR168222	AR168222 Sequence
23	94.6	93.7	592	6	AR177246	AR177246 Sequence
24	94.6	93.7	3308	6	CO818826	CO818826 Sequence
25	94.6	93.7	3308	14	AY334075	AY334075 Woodchuck
26	94.6	93.7	3308	14	AY334076	AY334076 Woodchuck
27	94.6	93.7	3308	14	AY334077	AY334077 Woodchuck
28	94.6	93.7	3308	14	AY628095	AY628095 Woodchuck
29	94.6	93.7	3308	14	AY628096	AY628096 Woodchuck
30	94.6	93.7	3308	14	AY628097	AY628097 Woodchuck
31	94.6	93.7	3308	14	AY628098	AY628098 Woodchuck
32	94.6	93.7	3308	14	AY628099	AY628099 Woodchuck
33	94.6	93.7	3308	14	AY628100	AY628100 Woodchuck
34	94.6	93.7	3308	14	OHVCGA	J02442 Woodchuck h
35	94.6	93.7	6893	6	AX823860	AX823860 Sequence
36	91.4	90.5	5617	6	AX384541	AX384541 Sequence
37	91.4	90.5	5691	6	AX359937	AX359937 Sequence
38	91.4	90.5	5691	6	AX382151	AX382151 Sequence
39	91.4	90.5	5711	6	AX359934	AX359934 Sequence
40	91.4	90.5	5711	6	AX382148	AX382148 Sequence
41	91.4	90.5	5731	6	AX384542	AX384542 Sequence
42	91.4	90.5	5732	6	AX359932	AX359932 Sequence
43	91.4	90.5	5732	6	AX382146	AX382146 Sequence
44	91.4	90.5	6026	6	AX384539	AX384539 Sequence
45	91.4	90.5	6140	6	AX384540	AX384540 Sequence

ALIGNMENTS

RESULT 1
AX937037
LOCUS AX937037 609 bp DNA linear PAT 06-JAN-2004
DEFINITION Sequence 1 from Patent EP1361277.
ACCESSION AX937037
VERSION AX937037.1 GI:40713229
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Mallet, J., Brun, S., Dufour, N. and Faucon-Biguet, N.
TITLE Optimization of transgene expression in mammalian cells
JOURNAL Patent: EP 1361277-A 1 12-NOV-2003;
CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS) (FR); Biovectys (FR)
FEATURES
source Location/Qualifiers
1..609
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Description of Artificial Sequence: WPRE region."

ORIGIN

Query Match 100.0%; Score 101.; DB 6; Length 609;
Best Local Similarity 100.0%; Pred. No. 1.2e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TGCCTTGCCTGCTGCTGACAGGGGCTCGGCTGTGGGCACTGACAAATTCGGTGGTGG 60
Db 338 TGCCTTGCCTGCTGCTGACAGGGGCTCGGCTGTGGGCACTGACAAATTCGGTGGTGG 397
Qy 61 TCGGGGAAGCTGACGCTCTTTCATGGCTGCTGCTGTGT 101
Db 398 TCGGGGAAGCTGACGCTCTTTCATGGCTGCTGCTGTGT 438
RESULT 2
AX952091
LOCUS AX952091 609 bp DNA linear PAT 08-JAN-2004
DEFINITION Sequence 1 from Patent WO03093485.
ACCESSION AX952091

Viruses; Retrovird viruses; Hepadnaviridae; Orthohepadnavirus.
1 (bases 1 to 2649)
Yamamoto, T., Litwin, S., Zhou, T., Zhu, Y., Condreay, L., Furman, P. and
Mason, W.S.

Mutations of the woodchuck hepatitis virus polymerase gene that
confer resistance to lamivudine and
2'-fluoro-5-methyl-beta-L-arabinofuranosyluracil
J. Virol. 76 (3), 1213-1223 (2002)

JOURNAL
MEDLINE
PUBMED
21635500
11773397

REFERENCE

2 (bases 1 to 2649)
Yamamoto, T., Litwin, S., Zhou, T., Zhu, Y., Condreay, L., Furman, P. and
Mason, W.S.

Direct Submission

Submitted (13-AUG-2001) Fox Chase Cancer Center, 7701 Burholme
Avenue, Philadelphia, PA 19111, USA
Location/Qualifiers

FEATURES

source
1. .2649
/organism="Woodchuck hepatitis B virus"
/virion
/mol_type="genomic DNA"
/isolate="342"
/db_xref="taxon:35269"
/note="isolated from serum of chronically infected
woodchuck"

CDS

1. .2649
/note="reverse transcriptase; contains six nucleotide
inframe deletion"
/codon_start=1
/product="defective polymerase"
/protein_id="AA04546.1"
/db_xref="GI:15637593"
/translation="MHFPSRLFRNIQSLGEEVEQLGPPEDALPLLAGEDLNHRVAD
ALNHLPTADLQWVHTNAITGLYSNOAOFNPHWIOPEPELHLHNDLQKLOQYFG
PLTNERKQLQNFAPRFPKATYFPLIKGINKNYENFALEHFFATANYLWLEAG
ILYLKRNQTLTLTKGKYSWEHRLQVHQHQKSHLQSRQNSMWACSGYLLHNLHP
SEPVSVTRLSNNISDKSOKSTRGLCSYKQIQTDRLHLARISCSGKITIGQGGSS
PKTSYKSSISNFRNQTWAYNSRNGHTTWFSSASNSKRSREKAYSSNSTKRYSP
NYEKSDSSFGVGRGRIITRLDNNGTTPQCLWRSFYNTKPCGSCYCTHHIVSSLDWGPCT
VTGVTYKSPTRPRITGGVFLVDKPNNSSESLVDFQSFGRHTRVHWKFAVEN
LQTLANLLSTNLQWLDVSAFVHIPISTAVPHLLVGSFGLERFTTCLSSHTNGN
DSQQTQWHTLCRTHIYSSLLLLFTYGRKHLHLAHPFMGRKLPNGVGLSPFLLAQF
TSALASWVRNPHCVFAYMDDLVLGARTSEHLTAIYSHICSVFLDLGHLNVRK
WMGNHLMFVGVITSSGVLQDQKHVKLSRYLSRVPNQPLDYKICERLTGILNYVAP
FTLCGIAALMPLHAIASRTAFIFSSLYKSWLLSLEYELWVVRQGVCTVADATP
TGMGIATTCOLLSTGTFAPLPATAEILAAACLRCTGACARLTGDNVSVLSGKLTSP
FWLLACVANWILRGTSFCVPSALNPADLPGRGLLPVLRPLRLRLPQTSRISLMAAS
PPVSPRRVRVAVSSPVQCEPWIPP"

ORIGIN

Query Match 100.0%; Score 101; DB 14; Length 2649;
Best Local Similarity 100.0%; Pred. No. 1.2e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TGCCTTCCCGCTGCTGACAGGGGCTCGGCTGTGGGCACTGACAAATTCGGTGGTGG 60
Db 2311 TGCCTTCCCGCTGCTGACAGGGGCTCGGCTGTGGGCACTGACAAATTCGGTGGTGG 2370
Qy 61 TCGGGGAAGCTGACGCTCTTCCATGCTGCTGCTGCTGTGT 101
Db 2371 TCGGGGAAGCTGACGCTCTTCCATGCTGCTGCTGCTGTGT 2411

RESULT 5

AF410857
LOCUS
DEFINITION Woodchuck hepatitis B virus isolate 335 type I mutant polymerase
gene, complete cds.
ACCESSION AF410857
VERSION AF410857.1 GI:15637590
KEYWORDS Woodchuck hepatitis B virus
SOURCE Woodchuck hepatitis B virus
ORGANISM Woodchuck hepatitis B virus

Viruses; Retrovird viruses; Hepadnaviridae; Orthohepadnavirus.
1 (bases 1 to 2655)
Yamamoto, T., Litwin, S., Zhou, T., Zhu, Y., Condreay, L., Furman, P. and
Mason, W.S.

Mutations of the woodchuck hepatitis virus polymerase gene that
confer resistance to lamivudine and
2'-fluoro-5-methyl-beta-L-arabinofuranosyluracil
J. Virol. 76 (3), 1213-1223 (2002)

JOURNAL
MEDLINE
PUBMED
21635500
11773397

REFERENCE

2 (bases 1 to 2655)
Yamamoto, T., Litwin, S., Zhou, T., Zhu, Y., Condreay, L., Furman, P. and
Mason, W.S.

Direct Submission

Submitted (13-AUG-2001) Fox Chase Cancer Center, 7701 Burholme
Avenue, Philadelphia, PA 19111, USA
Location/Qualifiers

FEATURES

source
1. .2655
/organism="Woodchuck hepatitis B virus"
/virion
/mol_type="genomic DNA"
/isolate="335"
/db_xref="taxon:35269"
/note="lamivudine resistant
isolated from serum of chronically infected woodchuck"

CDS

1. .2655
/note="reverse transcriptase"
/codon_start=1
/product="type I mutant polymerase"
/protein_id="AA04545.1"
/db_xref="GI:15637591"
/translation="MHFPSRLFRNIQSLGEEVEQLGPPEDALPLLAGEDLNHRVAD
ALNHLPTADLQWVHTNAITGLYSNOAOFNPHWIOPEPELHLHNDLQKLOQYFG
PLTNERKQLQNFAPRFPKATYFPLIKGINKNYENFALEHFFATANYLWLEAG
ILYLKRNQTLTLTKGKYSWEHRLQVHQHQKSHLQSRQNSMWACSGYLLHNLHP
SEPVSVTRLSNNISDKSOKSTRGLCSYKQIQTDRLHLARISCSGKITIGQGGSS
PKTSYKSSISNFRNQTWAYNSRNGHTTWFSSASNSKRSREKAYSSNSTKRYSP
CLNYEKSDSSFGVGRGRIITRLDNNGTTPQCLWRSFYNTKPCGSCYCTHHIVSSLDWGPCT
VTGVTYKSPTRPRITGGVFLVDKPNNSSESLVDFQSFGRHTRVHWKFAVEN
LQTLANLLSTNLQWLDVSAFVHIPISTAVPHLLVGSFGLERFTTCLSSHTNGN
DSQQTQWHTLCRTHIYSSLLLLFTYGRKHLHLAHPFMGRKLPNGVGLSPFLIT
QFTSALASWVRNPHCVFAYMDDLVLGARTSEHLTAIYSHICSVFLDLGHLNVRK
TKWGNHLMFVGVITSSGVLQDQKHVKLSRYLSRVPNQPLDYKICERLTGILNYVAP
FTLCGIAALMPLHAIASRTAFIFSSLYKSWLLSLEYELWVVRQGVCTVADATP
TGMGIATTCOLLSTGTFAPLPATAEILAAACLRCTGACARLTGDNVSVLSGKLTSP
FWLLACVANWILRGTSFCVPSALNPADLPGRGLLPVLRPLRLRLPQTSRISLMAAS
PPVSPRRVRVAVSSPVQCEPWIPP"

ORIGIN

Query Match 100.0%; Score 101; DB 14; Length 2655;
Best Local Similarity 100.0%; Pred. No. 1.2e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TGCCTTCCCGCTGCTGACAGGGGCTCGGCTGTGGGCACTGACAAATTCGGTGGTGG 60
Db 2317 TGCCTTCCCGCTGCTGACAGGGGCTCGGCTGTGGGCACTGACAAATTCGGTGGTGG 2376
Qy 61 TCGGGGAAGCTGACGCTCTTCCATGCTGCTGCTGCTGTGT 101
Db 2377 TCGGGGAAGCTGACGCTCTTCCATGCTGCTGCTGCTGTGT 2417

RESULT 6

AF410859
LOCUS
DEFINITION Woodchuck hepatitis B virus isolate 335 polymerase gene, complete
cds.
ACCESSION AF410859
VERSION AF410859.1 GI:15637594
KEYWORDS Woodchuck hepatitis B virus
SOURCE Woodchuck hepatitis B virus
ORGANISM Woodchuck hepatitis B virus

REFERENCE 1 (bases 1 to 2655)
AUTHORS Yamamoto,T., Litwin,S., Zhou,T., Zhu,Y., Condreay,L., Furman,P. and Mason,W.S.
TITLE Mutations of the woodchuck hepatitis virus polymerase gene that confer resistance to lamivudine and 2'-fluoro-5-methyl-beta-L-arabinofuranosyluracil
JOURNAL J. Virol. 76 (3), 1213-1223 (2002)
MEDLINE 21635500
PUBMED 11773397
REFERENCE 2 (bases 1 to 2655)
AUTHORS Yamamoto,T., Litwin,S., Zhou,T., Zhu,Y., Condreay,L., Furman,P. and Mason,W.S.
TITLE Direct Submission
JOURNAL Submitted (13-AUG-2001) Fox Chase Cancer Center, 7701 Burholme Avenue, Philadelphia, PA 19111, USA
FEATURES
source
1. .2655
/organism="Woodchuck hepatitis B virus"
/viroion
/mol_type="genomic DNA"
/isolate="335"
/db_xref="taxon:35269"
/note="isolated from serum of chronically infected woodchuck"
1. .2655
/note="reverse transcriptase"
/codon_start=1
/product="polymerase"
/protein_id="AAL04547.1"
/db_xref="GI:15637595"
/translation="MHPFSRLFRNIQSIGREGVQELLGPPEDALPLIAGBDLNRHVAD
ALNHLPTADLQVHWKTNATGLYSNQAQPNPHWIOPEPELHLDLHNDLQKLOQYRG
PLTINERKQLNFPARFPKATYFPLIKIGNNTPNFALEHFFATANYLTWLMEAG
ILYLKXNQTTLTFKGPYSWEHRQLVHNGQQHSHLQSQNSMWSACSGYLLHNLHP
SEPVSVSTRNLNNISDKSQSTRTGLCSYKQVQTDRLHLARI SCGSKITIGQOGSS
PKTSYKSISSNFRNOTWAYNSRNSGHTWFFSASNSKRSREKAYSSNSTKRYSP
PLNYEKSDPSSPGVGRGRTLDNNGTPQCLWBSFYNTKPGSYCIHHIVSSLDMDGP
CTVGDVTIKSPRTIRITGCVFLVDKNPNNSERSLVDFPSRGHTRVHWPKFAY
PNLQTLANLSTNLQWLSLDSAAFYHIPISPAAPVHLLVSGPGLERFTCLSSSTHN
GNDSQLQMTHTLCTRHLYSLLLFKTYGRKLHLAHPFMGRKLPFGMLGSPFLLA
QFTSALASVMVRNPHCVFAYMDLVLGARTSEHLTAIYSHICSVFLDLGILHNVK
TKWGNHLMFMGVYITSSGVLPODKVKLSRYLRSVPVNOPLDYKICERLTGLNVV
APFTLCGYAALMPLYHAIASRTAFIFSSLYKSWLLSLYELWVVRQGVVCTVFADA
TPTGWIATTCQLLSTGTFAPFLPIATABLIAACLRACWTGARLLGTDNSVVLGSKLTS
FPMWLLACVANWILRGTSFCYVPSALNPADLPSPRGLLPVLRLPLRLRPTQTSRLSMA
ASPPVSPRRPRVRVAWSSPVQTCPEWIPP"

ORIGIN
Query Match 100.0%; Score 101; DB 14; Length 2655;
Best Local Similarity 100.0%; Pred. No. 1.2e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAAATTCGGTGTGTG 60
Db 2317 TGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAAATTCGGTGTGTG 2376
Qy 61 TCGGGGAAGCTGACGCTCTTTCCATGGCTGCTCCGCTGTGT 101
Db 2377 TCGGGGAAGCTGACGCTCTTTCCATGGCTGCTCCGCTGTGT 2417

RESULT 7
AF410860
LOCUS
DEFINITION Woodchuck hepatitis B virus isolate 335 clone b polymerase gene, complete cds.
ACCESSION AF410860
VERSION AF410860.1 GI:15637596
KEYWORDS
SOURCE Woodchuck hepatitis B virus
ORGANISM Woodchuck hepatitis B virus
REFERENCE 1 (bases 1 to 2655)
Virus; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.

AUTHORS Yamamoto,T., Litwin,S., Zhou,T., Zhu,Y., Condreay,L., Furman,P. and Mason,W.S.
TITLE Mutations of the woodchuck hepatitis virus polymerase gene that confer resistance to lamivudine and 2'-fluoro-5-methyl-beta-L-arabinofuranosyluracil
JOURNAL J. Virol. 76 (3), 1213-1223 (2002)
MEDLINE 21635500
PUBMED 11773397
REFERENCE 2 (bases 1 to 2655)
AUTHORS Yamamoto,T., Litwin,S., Zhou,T., Zhu,Y., Condreay,L., Furman,P. and Mason,W.S.
TITLE Direct Submission
JOURNAL Submitted (13-AUG-2001) Fox Chase Cancer Center, 7701 Burholme Avenue, Philadelphia, PA 19111, USA
FEATURES
source
1. .2655
/organism="Woodchuck hepatitis B virus"
/viroion
/mol_type="genomic DNA"
/isolate="335"
/db_xref="taxon:35269"
/clone="b"
/note="isolated from serum of chronically infected woodchuck"
1. .2655
/note="reverse transcriptase"
/codon_start=1
/product="polymerase"
/protein_id="AAL04548.1"
/db_xref="GI:15637597"
/translation="MHPFSRLFRNIQSIGREGVQELLGPPEDALPLIAGBDLNRHVAD
ALNHLPTADLQVHWKTNATGLYSNQAQPNPHWIOPEPELHLDLHNDLQKLOQYRG
PLTINERKQLNFPARFPKATYFPLIKIGNNTPNFALEHFFATANYLTWLMEAG
ILYLKXNQTTLTFKGPYSWEHRQLVHNGQQHSHLQSQNSMWSACSGYLLHNLHP
SEPVSVSTRNLNNISDKSQSTRTGLCSYKQVQTDRLHLARI SCGSKITIGQOGSS
PKTSYKSISSNFRNOTWAYNSRNSGHTWFFSASNSKRSREKAYSSNSTKRYSP
PLNYEKSDPSSPGVGRGRTLDNNGTPQCLWBSFYNTKPGSYCIHHIVSSLDMDGP
CTVGDVTIKSPRTIRITGCVFLVDKNPNNSERSLVDFPSRGHTRVHWPKFAY
PNLQTLANLSTNLQWLSLDSAAFYHIPISPAAPVHLLVSGPGLERFTCLSSSTHN
GNDSQLQMTHTLCTRHLYSLLLFKTYGRKLHLAHPFMGRKLPFGMLGSPFLLA
QFTSALASVMVRNPHCVFAYMDLVLGARTSEHLTAIYSHICSVFLDLGILHNVK
TKWGNHLMFMGVYITSSGVLPODKVKLSRYLRSVPVNOPLDYKICERLTGLNVV
APFTLCGYAALMPLYHAIASRTAFIFSSLYKSWLLSLYELWVVRQGVVCTVFADA
TPTGWIATTCQLLSTGTFAPFLPIATABLIAACLRACWTGARLLGTDNSVVLGSKLTS
FPMWLLACVANWILRGTSFCYVPSALNPADLPSPRGLLPVLRLPLRLRPTQTSRLSMA
ASPPVSPRRPRVRVAWSSPVQTCPEWIPP"

ORIGIN
Query Match 100.0%; Score 101; DB 14; Length 2655;
Best Local Similarity 100.0%; Pred. No. 1.2e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAAATTCGGTGTGTG 60
Db 2317 TGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAAATTCGGTGTGTG 2376
Qy 61 TCGGGGAAGCTGACGCTCTTTCCATGGCTGCTCCGCTGTGT 101
Db 2377 TCGGGGAAGCTGACGCTCTTTCCATGGCTGCTCCGCTGTGT 2417

RESULT 8
AF410861
LOCUS
DEFINITION Woodchuck hepatitis B virus isolate 342 polymerase gene, complete cds.
ACCESSION AF410861
VERSION AF410861.1 GI:15637598
KEYWORDS
SOURCE Woodchuck hepatitis B virus
ORGANISM Woodchuck hepatitis B virus
REFERENCE 1 (bases 1 to 2655)
Virus; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.

/translation="MDIDPYKEFGSSVQLNPLDFFPDNLALVDTATLYEELTG
REHCSPHHTAIRQALVCWDELTKLIAMSSNITSEQVRILLVNHVNDTWGLKYRQSLM
FHLSCITFGHTVQBEFLVSFVWIRTPAPYRPNAPILSTLPEHTVIRRGAGARSRP
RRRTSPRRRSQSPRRRSQSPSANC"

ORIGIN

Query Match 100.0%; Score 101; DB 14; Length 3320;
Best Local Similarity 100.0%; Pred. No. 1.2e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGCCTTCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAAATCCGTTGGTGG 60
Db 1420 TGCCTTCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAAATCCGTTGGTGG 1479

Qy 61 TCGGGGAAGCTGAGCTCTTCCATGCTGCTCGCTGCTGTGT 101
Db 1480 TCGGGGAAGCTGAGCTCTTCCATGCTGCTCGCTGCTGTGT 1520

RESULT 10

OHVCGD 3323 bp ms-DNA circular VRL 04-MAY-1994
LOCUS Woodchuck hepatitis virus (WHV), complete genome, clone WHV 59.
DEFINITION M19183
ACCESSION M19183.1 GI:336141
VERSION DNA polymerase; WHC protein; WHSag protein; coat protein; complete genome; core protein; envelope-associated protein; origin of replication; polymerase.
KEYWORDS Woodchuck hepatitis B virus
SOURCE Woodchuck hepatitis B virus
ORGANISM Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
REFERENCE 1 (bases 1 to 3323)
AUTHORS Cohen, J.I., Miller, R.H., Rosenblum, B., Denniston, K., Gerin, J.L. and Purcell, R.H.
TITLE Sequence comparison of woodchuck hepatitis virus replicative forms
JOURNAL shows conservation of the genome
MEDLINE Virology 162 (1), 12-20 (1988)
PUBMED 88101359
COMMENT Original source text: Woodchuck hepatitis virus DNA, clone WHV 59, from a carrier woodchuck trapped in Pennsylvania.
FEATURES
source
1..3323
/organism="Woodchuck hepatitis B virus"
/mol_type="genomic DNA"
/db_xref="taxon:35269"
join(2427..3323,1..1758)
/codon_start=1
/product="DNA polymerase"
/protein_id="AAA46763.1"
/db_xref="GI:336143"
/translation="MHPFSRLFRNIQSLGEEVQELGPPEDALPLLAGBDLNRHVAD
ALNHLPTADLQWHTNAITGLYSNQAQFNPHWIOPEFELHNDLTKLQYVFG
PLTNEKRKLQNPAPFFPKATKYFPLIKIGNYPNFALEHFFATANYLWTWAG
ILYLKQNTLTTPKQKYSMEHRLQVHNGQKHLSHQNSWACSHLLHNLHP
SEPVSTRNLNSDKSQRKTGLCTKQVTDRLHARISGSKTIGQQGSS
PKTSYKISINFRNQWAYNSNSGHTWFSASNSKRSREKAYSNSSTSORVP
PLNKGDFSPGVRGRITRLDNNGLTPQCLMRSFYNTKPCGSIHHIYSLLDWWP
CTVTGDTIKSPRTPRTIGVLDKNNPNSSESLVDFVDFQPSRGRTHVHWPKAV
PNLQTLANLLTNLQSLDVSAAFYHIPISPAAPHLVGSFGLERFNTCMSSSTHN
GNSQLQTMALCTRHVYSSILLFLFTYGRKLHLAHFFIMGRKLPNGVGLSPFLIA
QFTSAISMVRRNPFPHCVFAYMDLVIGARTSEHLTAIYSHICSVFLDGLHNLVK
TKWGNHLHFMGYVITSSGLVPQDKVKLSRYLRSVPVQNPQDLYKICERLTGLNVY
APFTLCGYALPHTYAIASRTAFIFSSLYKSWLLSYELWLPVVRQGVVCTVFADA
TPTCWGIATTQCLSGTFAPPLPIATAELIAACIARCTGARLIGTNSVVLGSKLTS
PWLACVANWILRGTSFCVPSALNPADLPSCRLLPVLRPLRLRQPQTSRLSMA
ASPVSPRRPVRAWSPVQCEPWIP"

CDS

join(2992..3323,1..964)
/note="precursor"
/codon_start=1
/product="surface protein"
/protein_id="AAA46762.1"
/db_xref="GI:336142"

/translation="MGNNIKVFENPKIAAWPAVCTYTTTTPQNSVFPQGIYOTT
SLINPKNOELDSVLINRYKQIDMNTWQGFVVDOKLPLVSRDPLAPHLNOSQATFEI
KFGPIVFGIRDIPLGLVFPQPTNRDQGRKPTPTPLRDTHPHLTMKNOTRLOGF
VGLRLDITTEIRHNAIDPFITLSPVFTSVTILSPFTSGDPAISPEMSSSLGL
LAGLQVYFLMTIKILTIAQNLDMWMTSLSPFGIGPECTQNSQFQTKHLPTSCPTC
NGFRWMLRRFIYLLVLLCLIFLLVLDWKLIFVCPLOPTTETTVNCRQCTLSVQ
DTVTTPYCCCLAKTACNCTCWPISPSWALGNLYLMEWALARFSLNLLVPLLOWLGGIS
LTAWELLIMWFMGPAISILPPRIPIFVLFLIWWYI"
join(2992..3323,1..295)
/note="surface protein"
/product="surface protein"
296..961
1503..1928
/codon_start=1
/product="X protein"
/protein_id="AAA46764.1"
/db_xref="GI:336144"
/translation="NAARLCCQLDSARDVLLRPFQPSQSPFPFPPAAGSAASSTSS
PSPDSLDPLGLRLPACFASAGPCCLVFTCADLRTMDSTVNFVSWHAKRQLGWPSKD
LWPTYKDOLLTKWEGSIDPRLSIFVLGGRHKWRLL"
2021..2587
/codon_start=1
/product="core protein"
/protein_id="AAA46765.1"
/db_xref="GI:336145"
/translation="MDIDPYKEFGSSVQLNPLDFFPDNLALVDTATLYEELTG
REHCSPHHTAIRQALVCWDELTKLIAMSSNITSEQVRILLVNHVNDTWGLKYRQSLM
FHLSCITFGHTVQBEFLVSFVWIRTPAPYRPNAPILSTLPEHTVIRRGAGARSRP
RRRTSPRRRSQSPRRRSQSPSANC"

Query Match 100.0%; Score 101; DB 14; Length 3323;
Best Local Similarity 100.0%; Pred. No. 1.2e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGCCTTCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAAATCCGTTGGTGG 60
Db 1420 TGCCTTCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAAATCCGTTGGTGG 1479

Qy 61 TCGGGGAAGCTGAGCTCTTCCATGCTGCTCGCTGCTGTGT 101
Db 1480 TCGGGGAAGCTGAGCTCTTCCATGCTGCTCGCTGCTGTGT 1520

ORIGIN

RESULT 11

OHVCGD 3323 bp ms-DNA circular VRL 04-MAY-1994
LOCUS Woodchuck hepatitis virus (WHV), complete genome, clone WHV 7.
DEFINITION M18752
ACCESSION M18752
VERSION M18752.1 GI:336136
KEYWORDS DNA polymerase; coat protein; complete genome; core protein; envelope-associated protein; origin of replication; polymerase.
SOURCE Woodchuck hepatitis B virus
ORGANISM Woodchuck hepatitis B virus
REFERENCE 1 (bases 1 to 3323)
AUTHORS Cohen, J.I., Miller, R.H., Rosenblum, B., Denniston, K., Gerin, J.L. and Purcell, R.H.
TITLE Sequence comparison of woodchuck hepatitis virus replicative forms
JOURNAL shows conservation of the genome
MEDLINE Virology 162 (1), 12-20 (1988)
PUBMED 88101359
COMMENT Original source text: Woodchuck hepatitis virus DNA, clone WHV 7, from a carrier woodchuck trapped in Maryland.
FEATURES
source
1..3323
/organism="Woodchuck hepatitis B virus"
/mol_type="genomic DNA"
/db_xref="taxon:35269"
join(2427..3323,1..1758)
/codon_start=1
/product="DNA polymerase"

/protein_id="AAA46767.1"
/db_xref="GI:336138"
/translation="MHFSLFRNIQSLGEEVEQELGPPEDALPLLAGEDLNRHVD
ALNHLPTADLOWHKTNATIGLYSNOAOFNPHWIOPEFELHLHNDLLOKLOOYFG
PLTNEKRLQDLPNPAFFPKATKYVFLIKGINKNYENFALHFFATANYLWTLWEAG
ILYLRKQOTLTFGKQYSEHROLOVOHGOOKHSLQSRQNSWVACSGHLNHLIS
SESVSVTRNLNNISDKSQSTRIGLCSYKQIOTDRLEHARLSCSKYITIGQGS
PKLYLSSFNQWAINSSRNSGHTTWFSASNKNKREKAYLSTSKRSP
PLNYEKDFSGPVRRITRDLDNNGTPTQCLMRSFYNTKPGCYCIHIVSSLDWGP
CTVGDVTKISPRPTRTIGVFLVDKNPNNSSESLRVDFSQSRGHRTHVWPKFV
RNLQTLANLISLQWLSLVDVSAFYHPIPIPAAPVHLLVSGPLERFNTCLSSSTN
PNSOLOTMNLCTRHYVSSILLFKTVGRKLLAHFPMGPKLPMGVCISPLLA
OPTSALSMVRNRPCHVCFAYMDLVLGARTSEHLTAIYSHICSVELDLGHLNVK
TKWGNLHFMGYVITSGVLPQDKVKKISRYLRSVPVQPLDYKICERITGLTNY
APFTLCGYAALMPLYHATISRTAFISLSYKSWLLSYBELVVRGVCTVTFADA
TPTGCGATTYLLGFAPFLPIATIELAACLRACWTGARLIGTDSNVLSGLKTS
PFWLLACVANWILRGTSFCVVPVPSALNPADLPGLPVLRPLRLRQPOTSRLSIA
ASPVSPRPVRVAVSSPVQCEWIPP"
join(2992..3323,1..964)
/note="precursor"
/codon_start=1
/product="surface protein"
/protein_id="AAA46766.1"
/db_xref="GI:336137"
/translation="MGNINVTNPDKIAWPAVGYTYTTPQNSVFPQGIYQTT
SLINPKQQLDSVLINRYKQIDNTWQGVDPQKPLVSRDPPPKYINQSTHQTPEI
KPGPIIVGIRDIPIGLVPPQPTNRQGRKPTPTPLRTHPHTKMQTFHQFQF
LADGDLITTEORHNAVDPDTTSLSPAVPTVSTILSPSTTDDPALSPKSSLLQF
LAGIQQVYFLTKLTIAQNLDMWMTLSLSPGGIPECTGONSOFQCKHLPTSCPPNC
NGFRWYLRPIIYVLLVLLCLIFLLVLDWKGLIPVCPQPTTETTVNCRQCTISAQ
DMYTPYCCCLTAGNCTWPISSWALGNLWELARFSLNLLVPLQLQWGGIS
LIAWFLLIWIMFWGPAALLSILPFIPIFVFLIWIYI"
join(2992..3323,1..295)
/note="pre-surface protein"
296..961
/product="surface protein"
/codon_start=1
/product="X protein"
/protein_id="AAA46768.1"
/db_xref="GI:336139"
/translation="MAARLCCQLDSARDVLLRFPFGQSSGSPFPPPAAGSAASS
PSPSDSDLPGLRPLACFASAGPCCLVFTCAELRTMDSTVNFVSWHANRLQGMPSKD
LWTPYIKDQLLTKEEGSIDPRLSIFVLGGRHKCMRLI"
2021..2587
/codon_start=1
/product="core protein"
/protein_id="AAA46769.1"
/db_xref="GI:336140"
/translation="MDIDPYKEFGSSYQLLNFLPLDFFPDNLALVDATATAYEBELTG
REHCSPHHTAIROALVCWDELTKLIAMSSNITSEQVRTIIVNHVNDTWGLKVRQSLW
PHLSCLTFGQHTVQEFVLSFGVWIRTPAPYPPNAPILSTLPEHTVIRRGGAASRS
PRRTSPRRRSQSPRRRSQSPSANC"

CDS

sig_peptide

mat_peptide

CDS

1503..1928
/codon_start=1
/product="X protein"
/protein_id="AAA46768.1"
/db_xref="GI:336139"
/translation="MAARLCCQLDSARDVLLRFPFGQSSGSPFPPPAAGSAASS
PSPSDSDLPGLRPLACFASAGPCCLVFTCAELRTMDSTVNFVSWHANRLQGMPSKD
LWTPYIKDQLLTKEEGSIDPRLSIFVLGGRHKCMRLI"
2021..2587
/codon_start=1
/product="core protein"
/protein_id="AAA46769.1"
/db_xref="GI:336140"
/translation="MDIDPYKEFGSSYQLLNFLPLDFFPDNLALVDATATAYEBELTG
REHCSPHHTAIROALVCWDELTKLIAMSSNITSEQVRTIIVNHVNDTWGLKVRQSLW
PHLSCLTFGQHTVQEFVLSFGVWIRTPAPYPPNAPILSTLPEHTVIRRGGAASRS
PRRTSPRRRSQSPRRRSQSPSANC"

ORIGIN

Query Match 100.0%; Score 101; DB 14; Length 3323;
Best Local Similarity 100.0%; Pred. No. 1.2e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGCCTTCCCGCTGCTGACAGGGGCTCGGCTGTGGGCACTGACAAATCCGTCGTGTG 60
Db 1420 TGCCTTCCCGCTGCTGACAGGGGCTCGGCTGTGGGCACTGACAAATCCGTCGTGTG 1479

Qy 61 TCGGGGAAGCTGACGCTCTTCATGCTGCTCGCTGTGT 101
|||||

Db 1480 TCGGGGAAGCTGACGCTCTTCATGCTGCTCGCTGTGT 1520
|||||

RESULT 12

OHVHEPBA OHVHEPBA 3323 bp ss-DNA linear VRL 03-AUG-1993
LOCUS
DEFINITION Woodchuck hepatitis B virus (WHV8), complete genome.
ACCESSION J04514

J04514.1 GI:336146

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PubMed

COMMENT

FEATURES

source

Location/Qualifiers

1..3323

/organism="Woodchuck hepatitis B virus"

/mol_type="genomic DNA"

/db_xref="taxon:35269"

296..964

/note="surface protein"

/codon_start=1

/protein_id="AAA46770.1"

/db_xref="GI:336147"

/translation="MSPSPSLGLLAGLOVYVFLWTKLTIAQNLDMWMTLSLSPGGIP
ECTGONSOFQCKHLPTSCPPTCNGFRWYLRPIIYVLLVLDWKGLIPVCPQPTTETTVNCRQCTISAQDMYTPYCCCLTAGNCTWPISSWALGNLWELARFSLNLLVPLQLQWGGIS
LIAWFLLIWIMFWGPAALLSILPFIPIFVFLIWIYI"

WAVYI"

1503..1928

/note="X protein"

/codon_start=1

/protein_id="AAA46771.1"

/db_xref="GI:336148"

/translation="MAARLCCQLDSARDVLLRFPFGQSSGSPFPPPAAGSAASS
PSPSDSDLPGLRPLACFASAGPCCLVFTCAELRTMDSTVNFVSWHANRLQGMPSKD
LWTPYIKDQLLTKEEGSIDPRLSIFVLGGRHKCMRLI"

1719..1728

/note="direct repeat"

1941..1951

/note="direct repeat"

2021..2587

/note="core protein"

/codon_start=1

/protein_id="AAA46772.1"

/db_xref="GI:336149"

/translation="MDIDPYKEFGSSYQLLNFLPLDFFPDNLALVDATATAYEBELTG
REHCSPHHTAIROALVCWDELTKLIAMSSNITSEQVRTIIVNHVNDTWGLKVRQSLW
PHLSCLTFGQHTVQEFVLSFGVWIRTPAPYPPNAPILSTLPEHTVIRRGGAASRS
PRRTSPRRRSQSPRRRSQSPSANC"

PRRTSPRRRSQSPRRRSQSPSANC"

repeat_region

repeat_region

CDS

ORIGIN

Query Match 100.0%; Score 101; DB 14; Length 3323;
Best Local Similarity 100.0%; Pred. No. 1.2e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGCCTTCCCGCTGCTGACAGGGGCTCGGCTGTGGGCACTGACAAATCCGTCGTGTG 60
Db 1420 TGCCTTCCCGCTGCTGACAGGGGCTCGGCTGTGGGCACTGACAAATCCGTCGTGTG 1479

Qy 61 TCGGGGAAGCTGACGCTCTTCATGCTGCTCGCTGTGT 101
|||||

Db 1480 TCGGGGAAGCTGACGCTCTTCATGCTGCTCGCTGTGT 1520
|||||

RESULT 13

AY468486 AY468486 7207 bp DNA circular SYN 03-DEC-2003
LOCUS
DEFINITION Lentiviral transfer vector pHSCKW, complete sequence.
ACCESSION AY468486

VERSION AY468486.1 GI:38565537
KEYWORDS
SOURCE Lenticviral transfer vector pHSXW
ORGANISM Lenticviral transfer vector pHSXW
other sequences; artificial sequences; vectors.
REFERENCE 1 (bases 1 to 7207)
AUTHORS Johansen,J., Tornoe,J., Rosenblad,C., Dago,L. and Kusk,P.
TITLE Improved lenticviral transfer vector
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 7207)
AUTHORS Kusk,P. and Johansen,J.
TITLE Direct Submission
JOURNAL Submitted (19-NOV-2003) NsGene A/S, Baltorpvej 154, Ballerup DK-2750, Denmark
FEATURES
source
1. .7207
/organism="Lenticviral transfer vector pHSXW"
/mol_type="other DNA"
/db_xref="taxon:256321"
LTR 1. .634
/note="5' HIV-1 LTR"
misc_signal 686. .823
/note="packaging signal; psi"
misc_signal 1310. .1514
/note="Rev responsive element; RRE"
promoter 2031. .2547
/note="CMV promoter"
misc_feature 2618. .2671
/note="multiple cloning site"
misc_feature complement(2674. .2693)
/note="T7 recognition/binding site"
misc_feature 2711. .3315
/note="Woodchuck post-regulatory element; WPRE"
LTR 3512. .3746
/note="3' LTR; self-inactivating 3' LTR version; distal promoter of U3 deleted"
rep_origin 4719. .5392
gene complement(5537. .6397)
/gene="b1a"
CDS complement(5537. .6397)
/gene="b1a"
/note="ampicillin resistance gene"
/codon_start=1
/transl_table=11
/product="beta-lactamase"
/protein_id="AA24091.1"
/db_xref="GI:38565538"
/translation="MSIQHFRVALIPFFAFLPVPFAHPETLVKVKDAEDQLGARVGY
IELDLNSGKILSPFRPERFPMSTFKVLLCGAVLSRIDAGQQLGRIRHYSNDLVE
YSPVTEKHLDTGTVRELCSAALIMSENTAANLLLTIGGPKELTAPFLHNMGRHTL
DRWEPELNEALPNDERTTPVAVATTLRLKLLTGELLTLASRQOLIDWMEADKVGGLP
LRSLAPAGWFTADKSGAGERSGRIIAALGPDGKPSRIVVITYTGSQATMDERNRQIA
EIGASLIKHW"
ORIGIN
Query Match 100.0%; Score 101; DB 12; Length 7207;
Best Local Similarity 100.0%; Pred. No. 1.2e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TGCCTTCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAATTCGGTGGTGTG 60
Db 3048 TGCCTTCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAATTCGGTGGTGTG 3107
Qy 61 TCGGGGAAGCTGAGCTCTTTCCATGGCTGCTCGCCTGTGT 101
Db 3108 TCGGGGAAGCTGAGCTCTTTCCATGGCTGCTCGCCTGTGT 3148
RESULT 14
AX663053
LOCUS 7515 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 4 from Patent WO2006132.
ACCESSION AX663053

VERSION AX663053.1 GI:29163598
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE 1
AUTHORS Kingman,S.A., Mitrophanous,K. and Ellard,F.M.
TITLE Vector system
JOURNAL Patent: WO 02086132-A 4 31-OCT-2002;
Oxford Biomedica (UK) Limited (GB)
FEATURES
source
1. .7515
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="pSmart2 MCS 5prime cppt"
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 7515;
Best Local Similarity 100.0%; Pred. No. 1.2e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TGCCTTCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAATTCGGTGGTGTG 60
Db 4424 TGCCTTCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAATTCGGTGGTGTG 4483
Qy 61 TCGGGGAAGCTGAGCTCTTTCCATGGCTGCTCGCCTGTGT 101
Db 4484 TCGGGGAAGCTGAGCTCTTTCCATGGCTGCTCGCCTGTGT 4524
RESULT 15
MAR7
LOCUS 7808 bp DNA linear ROD 06-APR-1994
DEFINITION Woodchuck hepatitis virus surface antigen (pres) and X protein genes.
ACCESSION M60766
VERSION M60766.1 GI:191471
KEYWORDS X protein; surface protein.
SOURCE Marmota monax (woodchuck)
ORGANISM Marmota monax
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Scuridae; Sciurinae;
Marmota.
REFERENCE 1 (bases 1 to 7808)
AUTHORS Yamazoe,M., Nakai,S., Ogasawara,N. and Yoshikawa,H.
TITLE Integration of woodchuck hepatitis virus (WHV) DNA at two chromosomal sites (Vk and gag-like) in a hepatocellular carcinoma Gene 100, 139-146 (1991)
JOURNAL 91276235
MEDLINE
PUBMED 2055466
COMMENT source text: Woodchuck hepatitis virus DNA, clone B7.
FEATURES
Original Location/Qualifiers
source
1. .7808
/organism="Marmota monax"
/mol_type="genomic DNA"
/db_xref="taxon:9995"
/tissue_type="liver"
4123. .4496
/note="viral coding region"
/codon_start=3
/product="surface antigen"
/protein_id="AAA37108.1"
/db_xref="GI:191472"
CDS
TCWPIPSWALGNLWELARPSWLNLLVPLQLWLGISLIWFLLIWVFWGPAL
LSILPPIPIFLVFLIWIYI"
5035. .5505
/note="protein coding region crossing over viral/host boundary"
/codon_start=1
/protein_id="AAA37110.1"
/db_xref="GI:191474"
CDS

```

/translation="MAARLCCQLDSARDVLLRRFPQSSGPPPPRPAAGSAASSTSS
LSPDESDLPGLRPLPACFASAGPCCLVFTCADLRTMDSTVNFVSWHAKQLGMPKSD
LWTFYIKDQLLTDQLLTIDWMOYFKSPFLPSDFNMVSTPALDFMGLSLQFWI"
5035..5392
/note="viral coding region"
/codon_start=1
/product="X protein"
/protein_id="AAA37109.1"
/db_xref="GI:553849"
/translation="MAARLCCQLDSARDVLLRRFPQSSGPPPPRPAAGSAASSTSS
LSPDESDLPGLRPLPACFASAGPCCLVFTCADLRTMDSTVNFVSWHAKQLGMPKSD
LWTFYIKDQLLTDQLLT"
5590..5595

```

polyA_signal

ORIGIN

```

Query Match      100.0%; Score 101; DB 10; Length 7808;
Best Local Similarity 100.0%; Pred. No. 1.2e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  TGCCTTGCCCGCTGCTGGACAGGGGCTGGCTGTGGGCACTGACAATTCGCTGCTGTG 60
Db      4952 TGCCTTGCCCGCTGCTGGACAGGGGCTGGCTGTGGGCACTGACAATTCGCTGCTGTG 5011

Qy      61  TCGGGAGCTGACGTCCTTTCCATGGCTGCTCGCCTGTGT 101
Db      5012 TCGGGAGCTGACGTCCTTTCCATGGCTGCTCGCCTGTGT 5052

```

Search completed: July 14, 2005, 14:03:37
Job time : 757.618 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:35:42 ; Search time 142.398 Seconds
(without alignments)
4198.742 Million cell updates/sec

Title: US-09-482-682-65_COPY_4192_4292
Perfect score: 101
Sequence: 1 tgccttcccgcgtgctgac.....ccatggctgcgcctgtgt 101

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	592	9 ADA38373	Ada38373 Woodchuck
2	101	100.0	604	8 AAD55110	Aad55110 Woodchuck
3	101	100.0	604	8 ACC45094	Acc45094 Woodchuck
4	101	100.0	609	10 ADD67513	Add67513 WPRE elem
5	101	100.0	632	12 ADO25310	Ado25310 Woodchuck
6	101	100.0	2853	8 AAD55114	Aad55114 Hflapubig
7	101	100.0	2853	8 ACC45098	Acc45098 HIV-1 fla
8	101	100.0	7515	8 ABV77010	Abv77010 Nucleotid
9	101	100.0	7648	12 ADM47497	Adm47497 Lysosomal
10	101	100.0	8092	12 ADM47498	Adm47498 Lysosomal
11	101	100.0	8484	3 AAAS9091	Aaas9091 Nucleotid
12	101	100.0	8484	10 ADF48775	Adf48775 fibre exp
13	101	100.0	9731	11 ADM82791	Adm82791 DNA repai
14	101	100.0	9731	11 ADM82791	Adm82791 DNA repai
15	101	100.0	9782	11 ADM82792	Adm82792 DNA repai
16	101	100.0	10468	9 ACD27899	Acd27899 pdmT2 vec
17	94.6	93.7	592	2 AAX32299	Aax32299 Nucleotid
18	94.6	93.7	632	12 ADO01085	Ado01085 Woodchuck
19	94.6	93.7	6893	10 ADE24111	Ade24111 Proviral
20	91.4	90.5	3671	6 AAD28271	Aad28271 Alpha-lac

21	91.4	90.5	5617	6 AAD32077	Aad32077 Human alp
22	91.4	90.5	5691	6 AAD28313	Aad28313 Alpha-lac
23	91.4	90.5	5691	6 AAD28274	Aad28274 Alpha-lac
24	91.4	90.5	5691	12 ADM68974	Adm68974 Alpha-lac
25	91.4	90.5	5711	6 AAD28310	Aad28310 Alpha-lac
26	91.4	90.5	5711	12 ADM68971	Adm68971 Alpha-lac
27	91.4	90.5	5731	6 AAD32078	Aad32078 Human alp
28	91.4	90.5	5732	6 AAD28308	Aad28308 Mouse mam
29	91.4	90.5	5732	6 AAD28269	Aad28269 Mouse mam
30	91.4	90.5	5732	12 ADM68969	Adm68969 MMTV MN14
31	91.4	90.5	6026	6 AAD32075	Aad32075 Human alp
32	91.4	90.5	6027	12 ADL35206	Adl35206 Plasmid p
33	91.4	90.5	6140	6 AAD32076	Aad32076 Human alp
34	91.4	90.5	6706	12 ADL35208	Adl35208 Plasmid p
35	91.4	90.5	6748	12 ADL35207	Adl35207 Plasmid p
36	91.4	90.5	7033	12 ADO07393	Ado07393 Modified
37	91.4	90.5	7248	12 ADL35209	Adl35209 Plasmid p
38	91.4	90.5	7350	12 ADL35212	Adl35212 Plasmid p
39	91.4	90.5	7650	12 ADL35213	Adl35213 Plasmid p
40	91.4	90.5	7927	12 ADL35211	Adl35211 Plasmid p
41	91.4	90.5	7969	12 ADL35210	Adl35210 Plasmid p
42	91.4	90.5	9183	6 AAD28309	Aad28309 Alpha-lac
43	91.4	90.5	9183	6 AAD28270	Aad28270 Alpha-lac
44	91.4	90.5	9183	12 ADM68970	Adm68970 Alpha-lac
45	91.4	90.5	9291	10 ACH00963	Ach00963 PSMA anti

ALIGNMENTS

RESULT 1

ADA38373
ID - ADA38373 standard; DNA; 592 BP.

XX ADA38373;

DT 20-NOV-2003 (first entry)

XX Woodchuck hepatitis virus postranscriptional regulatory element.

XX woodchuck hepatitis virus; postranscriptional regulatory element; ds;

XX trans-lenti viral vector system; Vpr; Vpx; reverse transcriptase;

XX integrase; gag; gag-pol precursor;

XX polypurine tract-central terminator sequence; PPT-CTS; gene transfer;

XX gene expression; transduction; replication competent retrovirus; RCR.

XX Woodchuck hepatitis B virus.

XX US2003072938-A1.

XX 17-APR-2003.

XX 24-JUL-2002; 2002US-00202457.

XX 03-JUN-1998; 98US-00089900.

XX 14-DEC-1999; 99US-00460548.

XX 13-NOV-2000; 2000US-00709501.

XX (KAPP/) KAPPES J C.

XX (WUX/) WU X.

XX Kappes JC, Wu X;

XX WPI; 2003-625514/59.

XX New trans-lenti viral vector systems with cis-acting sequences.

XX polypurine tract-central terminators, etc., useful for gene transfer,

XX e.g. delivering genes encoding therapeutic or viral inhibitory

XX polypeptides to cells.

XX Example 5; Page 7; 19pp; English.

XX The invention describes a trans-lenti viral vector system. The new trans-

CC lenti viral vector system comprises: a first nucleic acid segment
CC encoding at least one fusion protein comprising a functional portion of a
CC Vpr or Vpx polypeptide fused in frame to a functional portion of a
CC Reverse transcriptase polypeptide fused in frame to a functional portion of a
CC of an integrase polypeptide, where the first nucleic acid segment is
CC capable of expression in a mammalian cell, and where the functional
CC portion of the Vpr or Vpx polypeptide is capable of providing for the
CC incorporation of the fusion protein into a viral particle; a second
CC nucleic acid segment comprising a nucleotide sequence encoding a
CC functional portion of a Gag polypeptide and functional portion of a
CC Protease polypeptide, where the second nucleic acid is capable of
CC expression in the mammalian, and where the second nucleic acid does not
CC encode a functional Reverse Transcriptase polypeptide or a functional
CC Integrase polypeptide; a third nucleic acid segment comprising a nucleic
CC acid sequence encoding a functional envelope polypeptide capable of
CC mediating recognition and entry of the viral particle into a target cell,
CC where the third nucleic acid segment does not encode a functional Gag-Pol
CC precursor; and a fourth nucleic acid segment comprising a heterologous
CC nucleic acid sequence and at least one nucleotide sequence consisting of
CC a functional equivalent of a polypurine tract-central terminator sequence
CC (PPT-CTS), a functional equivalent of Woodchuck Hepatitis Virus Post
CC transcriptional Regulatory Element (WPRE), a PPT-CTS or a WPRE. The trans
CC -viral vector system produces a viral particle capable of introducing the
CC heterologous nucleotide sequence into the genome of the target cell. The
CC target cell is a non-dividing cell, a primary cell, macrophage, a CD34+
CC cell, or haematopoietic stem cell. The trans-lenti viral vector system is
CC useful for gene transfer, particularly for producing viruses or viral
CC particles capable of introducing heterologous nucleotide sequence(s) into
CC the genome of a target cell. The trans-lenti viral vector system is
CC particularly useful for delivering heterologous nucleic acid sequences
CC (which encode viral inhibitory polypeptides or a therapeutic
CC polypeptides) to cells with greater efficiency and/or effect, and for
CC improving gene expression. The present lentiviral vector system affords
CC relatively high vector particle production, has improved transduction
CC capabilities, and has even lower potential for replication competent
CC retrovirus (RCR) events. This sequence represents woodchuck hepatitis
CC virus posttranscriptional regulatory element (WPRE) used in the creation
CC of the trans-lenti viral vector system.

XX SQ Sequence 592 BP; 77 A; 188 C; 148 G; 179 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 9; Length 592;
Best Local Similarity 100.0%; Pred. No. 4.8e-22;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGCCTTCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCGGTGGTTG 60
Db 328 TGCCTTCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCGGTGGTTG 387

Qy 61 TCGGGGAAGCTGACGCTCTTCCATGGCTGCTGCCCTGTGT 101
Db 388 TCGGGGAAGCTGACGCTCTTCCATGGCTGCTGCCCTGTGT 428

RESULT 2

AAD55110
ID AAD55110 standard; DNA; 604 BP.

XX AC AAD55110;
XX XX
XX 27-OCT-2003 (revised)
DT 07-AUG-2003 (first entry)

XX Woodchuck hepatitis virus regulator element (WRE) DNA.

XX Transgenic; biotechnology; Woodchuck hepatitis virus regulator element;
XX WRE; ds.

XX Woodchuck hepatitis B virus.

XX WO2003022040-A2.

XX 20-MAR-2003.

XX 12-SEP-2002; 2002WO-US029130.
XX 13-SEP-2001; 2001US-0322031P.
PR 09-JAN-2002; 2002US-0347782P.
XX (CALY) CALIFORNIA INST OF TECHNOLOGY.

XX Baltimore D, Hong EJ, Lois-Caballe C, Pease S;
PI WPI; 2003-300976/29.

XX Producing a transgenic animal for commercial use, comprises transfecting
PT a packaging cell line with retroviral construct, recovering recombinant
PT retrovirus from the cell line and infecting embryonic cell with the
PT recombinant virus.

XX Example 1; Fig 18B; 76pp; English.

XX The invention relates to a method for producing a transgenic animal for
CC commercial use, which comprises transfecting a packaging cell line with
CC retroviral construct, recovering recombinant retrovirus from the cell
CC line and infecting embryonic cell with the recombinant virus. The method
CC is useful in producing transgenic animals using retroviral constructs
CC engineered to carry a transgene of interest. The transgenic animals may
CC find use in commercial applications like biotechnology and agriculture.
CC The present sequence is Woodchuck hepatitis virus regulator element (WRE)
CC DNA used to illustrate the method of the invention. (Updated on 27-OCT-
CC 2003 to standardise OS field)

XX SQ Sequence 604 BP; 81 A; 190 C; 150 G; 183 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 8; Length 604;
Best Local Similarity 100.0%; Pred. No. 4.8e-22;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TGCCTTCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCGGTGGTTG 60
Db 334 TGCCTTCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCGGTGGTTG 393

Qy 61 TCGGGGAAGCTGACGCTCTTCCATGGCTGCTGCCCTGTGT 101
Db 394 TCGGGGAAGCTGACGCTCTTCCATGGCTGCTGCCCTGTGT 434

RESULT 3
ACC45094
ID ACC45094 standard; DNA; 604 BP.

XX AC ACC45094;

XX 27-OCT-2003 (revised)
DT 10-JUN-2003 (first entry)

XX Woodchuck hepatitis virus regulator element (WRE) SEQ ID NO:4.

XX Transgenic animal; transgenic bird; transgenic fish; transgene;
KW retroviral construct; lentiviral; long terminal repeat; LTR; WRE;
KW biotechnology; agriculture; woodchuck hepatitis virus; regulator element;
KW gene; ds.

XX Woodchuck hepatitis B virus.

XX WO2003022228-A2.

XX 20-MAR-2003.

XX 12-SEP-2002; 2002WO-US029157.

XX 13-SEP-2001; 2001US-0322031P.

PR 09-JAN-2002; 2002US-0347782P.

XX (CALY) CALIFORNIA INST OF TECHNOLOGY.


```

XX Baltimore D, Hong EJ, Lois-Caballe C, Pease S;
XX WPI; 2003-301005/29.
XX Producing a transgenic bird or fish for commercial use, comprises
XX transfecting a packaging cell line with retroviral construct, recovering
XX recombinant retrovirus from the cell and infecting bird or fish egg with
XX the recombinant virus.
XX
XX Example 1; Fig 18; 68pp; English.
XX
XX The present invention describes a method for producing a transgenic bird
XX or fish. The method comprises transfecting a packaging cell line with a
XX retroviral construct, recovering recombinant retroviral particles from
XX the packaging cell line, and infecting a bird or a fish egg with the
XX recombinant retroviral particles. The retroviral construct comprises the
XX R and U5 sequences from a 5' lentiviral long terminal repeat (LTR) and a
XX self-inactivating 3' lentiviral LTR. Also described is a transgenic bird
XX or fish made by the above method and whose genome comprises a proviral
XX DNA that has a self-inactivating 3' lentiviral LTR. The method is useful
XX in producing transgenic animals, particularly transgenic birds and fish,
XX using retroviral constructs engineered to carry a transgene of interest.
XX The method is used to introduce the gene of choice into animals in order
XX to confer upon them desired attributes. The transgenic animals may find
XX use in commercial applications like biotechnology and agriculture. The
XX present sequence represents a woodchuck hepatitis virus regulator element
XX (WRE) nucleotide sequence, which is used in an example from the present
XX invention. (Updated on 27-OCT-2003 to standardise OS field)
XX
XX Sequence 604 BP; 81 A; 190 C; 150 G; 183 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 101; DB 8; Length 604;
XX Best Local Similarity 100.0%; Pred. No. 4.8e-22;
XX Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 1 TGCCTTCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAAATTCGGTGTGTTG 60
XX Db 334 TGCCTTCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAAATTCGGTGTGTTG 393
XX
XX Qy 61 TCGGGGAAGCTGACGTCTCTTCCATGGCTGCTCGCCTGTGT 101
XX Db 394 TCGGGGAAGCTGACGTCTCTTCCATGGCTGCTCGCCTGTGT 434
XX
XX RESULT 4
XX ADD67513
XX ID ADD67513 standard; DNA; 609 BP.
XX AC ADD67513;
XX
XX DT 29-JAN-2004 (first entry)
XX DE WPRE element #SEQ ID 1.
XX
XX Neuroprotective; antiparkinsonian; neurotropic; anticonvulsant;
XX transgene delivery; WPRE element; APP5'UTR; tau3'UTR; TH3'UTR; vector;
XX neurodegenerative disease; Parkinson's disease; Alzheimer's disease;
XX amyotrophic lateral sclerosis; Huntington's disease;
XX retinal degenerative disease; posttranscription; ds.
XX
XX OS Synthetic.
XX
XX PN EP1361277-A1.
XX
XX PD 12-NOV-2003.
XX
XX 30-APR-2002; 2002EP-00291091.
XX
XX 30-APR-2002; 2002EP-00291091.
XX
XX (CNRS ) CNRS CENT NAT RECH SCI.
XX (BIOV-) BIOVECTYS.
XX
XX Baltimore D, Brun S, Dufour N, Faucon-Biguet N;
XX WPI; 2003-879907/82.
XX
XX New vector, for transgene delivery into mammalian cells, comprising a
XX chimeric genetic construct with a transgene linked to a WPRE element, or
XX APP5'UTR, tau3'UTR or TH3'UTR region, useful for treating
XX neurodegenerative disease.
XX
XX Claim 9; SEQ ID NO 1; 30pp; English.
XX
XX The invention relates to a vector for transgene delivery into mammalian
XX cells. The vector comprises a chimeric genetic construct with a transgene
XX operably linked to at least two distinct posttranscriptional regulatory
XX elements, e.g. WPRE element, APP5'UTR, tau3'UTR or TH3'UTR region. The
XX WPRE element, APP5'UTR, tau3'UTR and TH3'UTR region or their functional
XX fragment comprises the nucleotide sequence of 609, 95, 237 and 91 bp,
XX respectively ADD67513-ADD67516. The vector comprises a promoter
XX controlling transcription of the transgene in the mammalian cells, a
XX marker gene and a polyadenylation signal operably linked to the
XX transgene. The vector is a plasmid or a recombinant virus. The vector or
XX recombinant cell is used for the manufacture of a medicament for treating
XX a disease. The vector, recombinant cell or composition is useful for
XX treating a human disease, e.g. neurodegenerative diseases selected from
XX Parkinson's disease, Alzheimer's disease, amyotrophic lateral sclerosis,
XX Huntington's disease or retinal degenerative diseases. They can also be
XX used in experiments research or prophylactic areas. The current sequence
XX represents the WPRE element nucleotide sequence.
XX
XX Sequence 609 BP; 83 A; 191 C; 151 G; 184 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 101; DB 10; Length 609;
XX Best Local Similarity 100.0%; Pred. No. 4.8e-22;
XX Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 1 TGCCTTCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAAATTCGGTGTGTTG 60
XX Db 338 TGCCTTCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAAATTCGGTGTGTTG 397
XX
XX Qy 61 TCGGGGAAGCTGACGTCTCTTCCATGGCTGCTCGCCTGTGT 101
XX Db 398 TCGGGGAAGCTGACGTCTCTTCCATGGCTGCTCGCCTGTGT 438
XX
XX RESULT 5
XX ADO25310
XX ID ADO25310 standard; DNA; 632 BP.
XX AC ADO25310;
XX
XX DT 12-AUG-2004 (first entry)
XX
XX Woodchuck hepatitis virus WPRE sequence as biological cloning marker.
XX
XX ss; porcine; uroplakin II gene; promoter; expression vector;
XX surrogate mother animal; transgenic animal; urine; bladder.
XX
XX OS Woodchuck hepatitis B virus.
XX
XX PN WC2004042062-A1.
XX
XX PD 21-MAY-2004.
XX
XX 04-NOV-2003; 2003WO-KR002339.
XX
XX 04-NOV-2002; 2002KR-00067856.
XX
XX 03-NOV-2003; 2003KR-00077256.
XX
XX (CHOA-) CHO-A PHARM CO LTD.
XX (KIMJ/) KIM J.
XX
XX Kim J;

```

XX WPI; 2004-411520/38.
 XX
 XX Novel porcine uroplakin II gene promoter, useful for promoting the
 XX bladder-specific expression of a specific target protein.
 XX
 XX Claim 7; SEQ ID NO 7; 76pp; English.
 XX
 XX The invention relates to a novel porcine uroplakin II gene promoter (I).
 XX An expression vector containing the promoter is useful in a method of
 XX producing useful proteins which involves implanting the vector into a
 XX surrogate mother animal, obtaining transgenic animals from the surrogate
 XX mother animal, and isolating and purifying useful proteins from the urine
 XX of the transgenic animals. The promoter, the expression vector, and the
 XX transgenic animal can be used in the production field of useful proteins
 XX that are medically valuable. The promoter promotes the bladder-specific
 XX expression of a target protein at high efficiency. An animal which was
 XX transformed using the promoter, so as to express the target protein,
 XX secret with a target protein in its urine at high concentration, and the
 XX protein thus produced shows a superior physiological activity to that of
 XX the same kind of the existing protein. This sequence represents the
 XX woodchuck hepatitis virus posttranslational regulatory element (WPRE)
 XX sequence which is used as a marker gene for cloning of genes of interest
 XX and linking to the porcine uroplakin II gene promoter.
 XX
 XX Sequence 632 BP; 80 A; 201 C; 158 G; 193 T; 0 U; 0 Other;
 XX
 XX Query Match 100.0%; Score 101; DB 12; Length 632;
 XX Best Local Similarity 100.0%; Pred. No. 4.9e-22;
 XX Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 1 TGCTTCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAAATCCGTTGTTG 60
 XX Db 347 TGCTTCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAAATCCGTTGTTG 406
 XX
 XX QY 61 TCGGGGAAGCTGACGCTCTTTCATGGCTGCTGCCCTGTGT 101
 XX Db 407 TCGGGGAAGCTGACGCTCTTTCATGGCTGCTGCCCTGTGT 447
 XX
 XX RESULT 6
 XX AAD55114
 XX ID AAD55114 standard; DNA; 2853 BP.
 XX AC AAD55114;
 XX
 XX DT 07-AUG-2003 (first entry)
 XX
 XX DE HflapUbiGWRE chimeric construct DNA.
 XX
 XX KW Transgenic; biotechnology; agriculture; green fluorescent protein; GFP;
 XX KW Woodchuck hepatitis virus regulator element; WRE; human; ubiquitin;
 XX KW Human immunodeficiency virus type 1; HIV-1; chimeric; ds.
 XX
 XX OS Human immunodeficiency virus 1.
 XX OS Unidentified.
 XX OS Homo sapiens.
 XX OS Woodchuck hepatitis B virus.
 XX OS Chimeric.
 XX
 XX PN WO2003022040-A2.
 XX
 XX PD 20-MAR-2003.
 XX
 XX PF 12-SEP-2002; 2002WO-US029130.
 XX
 XX PR 13-SEP-2001; 2001US-0322031P.
 XX PR 09-JAN-2002; 2002US-0347782P.
 XX
 XX PA (CALY) CALIFORNIA INST OF TECHNOLOGY.
 XX
 XX PI Baltimore D, Hong EJ, Lois-Caballe C, Pease S;
 XX
 XX WPI; 2003-301005/29.
 XX
 XX Producing a transgenic bird or fish for commercial use, comprises
 XX transfecting a packaging cell line with retroviral construct, recovering

DR WPI; 2003-300976/29.
 XX
 XX Producing a transgenic animal for commercial use, comprises transfecting
 XX a packaging cell line with retroviral construct, recovering recombinant
 XX retrovirus from the cell line and infecting embryonic cell with the
 XX recombinant virus.
 XX
 XX Example 1; Fig 21; 76pp; English.
 XX
 XX The invention relates to a method for producing a transgenic animal for
 XX commercial use, which comprises transfecting a packaging cell line with
 XX retroviral construct, recovering recombinant retrovirus from the cell
 XX line and infecting embryonic cell with the recombinant virus. The method
 XX is useful in producing transgenic animals using retroviral constructs
 XX engineered to carry a transgene of interest. The transgenic animals may
 XX find use in commercial applications like biotechnology and agriculture.
 XX The present sequence is HflapUbiGWRE chimeric construct DNA comprising
 XX Human immunodeficiency virus type 1 (HIV-1) NL4.3 flap sequence, green
 XX fluorescent protein (GFP) variant encoding sequence, human ubiquitin
 XX promoter sequence and Woodchuck hepatitis virus regulator element (WRE).
 XX This sequence is used to illustrate the method of the invention
 XX
 XX Sequence 2853 BP; 564 A; 774 C; 858 G; 657 T; 0 U; 0 Other;
 XX
 XX Query Match 100.0%; Score 101; DB 8; Length 2853;
 XX Best Local Similarity 100.0%; Pred. No. 6.5e-22;
 XX Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 1 TGCTTCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAAATCCGTTGTTG 60
 XX Db 2559 TGCTTCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAAATCCGTTGTTG 2618
 XX
 XX QY 61 TCGGGGAAGCTGACGCTCTTTCATGGCTGCTGCCCTGTGT 101
 XX Db 2619 TCGGGGAAGCTGACGCTCTTTCATGGCTGCTGCCCTGTGT 2659
 XX
 XX RESULT 7
 XX ACC45098
 XX ID ACC45098 standard; DNA; 2853 BP.
 XX AC ACC45098;
 XX
 XX DT 10-JUN-2003 (first entry)
 XX
 XX DE HIV-1 flap + ubiquitin + GFP + WRE construct DNA sequence SEQ ID NO:8.
 XX
 XX KW Transgenic animal; transgenic bird; transgenic fish; transgene;
 XX KW retroviral construct; lentiviral; long terminal repeat; LTR;
 XX KW biotechnology; agriculture; gene; ds.
 XX
 XX OS Human immunodeficiency virus 1.
 XX OS Homo sapiens.
 XX OS Woodchuck hepatitis B virus.
 XX OS Synthetic.
 XX
 XX PN WO2003022228-A2.
 XX
 XX PD 20-MAR-2003.
 XX
 XX PF 12-SEP-2002; 2002WO-US029157.
 XX
 XX PR 13-SEP-2001; 2001US-0322031P.
 XX PR 09-JAN-2002; 2002US-0347782P.
 XX
 XX PA (CALY) CALIFORNIA INST OF TECHNOLOGY.
 XX
 XX PI Baltimore D, Hong EJ, Lois-Caballe C, Pease S;
 XX
 XX WPI; 2003-301005/29.
 XX
 XX Producing a transgenic bird or fish for commercial use, comprises
 XX transfecting a packaging cell line with retroviral construct, recovering

PT recombinant retrovirus from the cell and infecting bird or fish egg with
XX the recombinant virus.
XX
XX Example 1; Fig 21; 68pp; English.
XX
CC The present invention describes a method for producing a transgenic bird
CC or fish. The method comprises transfecting a packaging cell line with a
CC retroviral construct, recovering recombinant retroviral particles from
CC the packaging cell line, and infecting a bird or a fish egg with the
CC recombinant retroviral particles. The retroviral construct comprises the
CC R and U5 sequences from a 5' lentiviral long terminal repeat (LTR) and a
CC self-inactivating 3' lentiviral LTR. Also described is a transgenic bird
CC or fish made by the above method and whose genome comprises a proviral
CC DNA that has a self-inactivating 3' lentiviral LTR. The method is useful
CC in producing transgenic animals, particularly transgenic birds and fish,
CC using retroviral constructs engineered to carry a transgene of interest.
CC The method is used to introduce the gene of choice into animals in order
CC to confer upon them desired attributes. The transgenic animals may find
CC use in commercial applications like biotechnology and agriculture. The
CC present sequence represents a construct nucleotide sequence comprising an
CC HIV-1 sequence, a green fluorescent protein (GFP) variant sequence, a
CC human ubiquitin promoter sequence and a woodchuck hepatitis regulator
CC element sequence, which is used in an example from the present invention
XX
SQ Sequence 2853 BP; 564 A; 774 C; 858 G; 657 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 8; Length 2853;
Best Local Similarity 100.0%; Pred. No. 6.5e-22;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTGGCACTGACAATTCGGTGTGTTG 60
Db 2559 TGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTGGCACTGACAATTCGGTGTGTTG 2618

QY 61 TCGGGGAAGCTGACGCTCTTTCCATGGCTGCTCGCCTGTGT 101
Db 2619 TCGGGGAAGCTGACGCTCTTTCCATGGCTGCTCGCCTGTGT 2659

RESULT 8
ABV77010
ID ABV77010 standard; DNA; 7515 BP.
XX
XX ABV77010;
AC
XX
XX 03-MAR-2003 (first entry)
DT
XX
XX Nucleotide sequence of plasmid pSmart2 5'cppt.
DE
XX
XX Viral vector; adipose tissue; adipose tissue metabolism; obesity;
KW diabetes; blood disorder; vascular disease; ss.
XX
XX Synthetic.
OS
XX
XX WO200286132-A2.
FN
XX
XX 31-OCT-2002.
PD
XX
XX 19-APR-2002; 2002WO-GB001830.
PP
XX
XX 20-APR-2001; 2001GB-00009781.
PR
XX
XX (OXFO-) OXFORD BIOMEDICA UK LTD.
PA
XX
XX Kingsman SA, Mitrophanous K, Ellard FM;
PI
XX
XX WPI; 2003-093139/08.
DR
XX
XX Use of viral vector system for transducing a target adipose tissue site,
PT and for treating and/or preventing vascular diseases or diseases
PT associated with death or impaired function of adipose tissue cells, such
PT as obesity and diabetes.
XX

PS Example; Page 65-67; 84pp; English.
XX
XX The specification describes the use of a viral vector system for
CC transducing a target adipose tissue site. The viral vector system is
CC useful for transducing a target adipose tissue site, in the manufacture
CC of a pharmaceutical composition for treating and/or preventing a disease
CC associated with a derangement in the metabolism of adipose tissue, such
CC as obesity and diabetes. The viral vector system is also useful for
CC treating and preventing a disease associated with death or impaired
CC function of adipose tissue cells, a disease associated with hereditary
CC blood disorders, and vascular diseases. The present sequence represents
CC plasmid pSmart 5'cppt, which is used to construct vectors for use in the
CC invention
XX
SQ Sequence 7515 BP; 1978 A; 1733 C; 1829 G; 1975 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 8; Length 7515;
Best Local Similarity 100.0%; Pred. No. 7.8e-22;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCTTTCGCCGCTGCTGGACAGGGGCTCGGCTGTGGCACTGACAATTCGGTGTGTTG 60
Db 4424 TGCTTTCGCCGCTGCTGGACAGGGGCTCGGCTGTGGCACTGACAATTCGGTGTGTTG 4483

QY 61 TCGGGGAAGCTGACGCTCTTTCCATGGCTGCTCGCCTGTGT 101
Db 4484 TCGGGGAAGCTGACGCTCTTTCCATGGCTGCTCGCCTGTGT 4524

RESULT 9
ADM47497
ID ADM47497 standard; DNA; 7648 BP.
XX
XX ADM47497;
AC
XX
XX 03-JUN-2004 (first entry)
DT
XX
XX Lysosomal enzyme related DNA #1.
DE
XX
XX Lysosomal enzyme; lysosomal storage disease; Gaucher type I disease;
KW Hurler disease; Sanfilippo disease; gene therapy; ds.
XX
XX Unidentified.
OS
XX
XX US2004023218-A1.
FN
XX
XX 05-FEB-2004.
PD
XX
XX 21-JUN-2002; 2002US-00176066.
PP
XX
XX 21-JUN-2002; 2002US-00176066.
PR
XX
XX (DESM/) DESMARIS N.
PA
XX
XX (HEAR/) HEARD J M.
PA
XX
XX Desmaris N, Heard JM;
PI
XX
XX WPI; 2004-142648/14.
DR
XX
XX New purified nucleic acid molecules capable of expressing a lysosomal
PT enzyme, useful for preventing or treating lysosomal storage diseases
PT (e.g. Gaucher type I disease or Sanfilippo disease) in humans.
XX
XX Disclosure; Page 9-11; 20pp; English.
XX
XX The invention relates to a nucleic acid molecule capable of expressing a
CC lysosomal enzyme. The nucleic acid molecule comprises at least a sequence
CC coding for the lysosomal enzyme and a promoter highly active in the brain
CC inserted upstream from the sequence. Compositions and methods of the
CC invention are useful for preventing or treating lysosomal storage
CC diseases e.g. Gaucher type I disease, Hurler disease (MPSI) or Sanfilippo
CC disease (MPSIII) in humans. The invention is also useful in gene therapy.
CC The present sequence is a DNA related to the invention.
XX

XX SQ Sequence 7648 BP; 1546 A; 2244 C; 2052 G; 1806 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 12; Length 7648;
Best Local Similarity 100.0%; Pred. No. 7.9e-22;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAATTCGGTGGTGTG 60
|||||
Db 3303 TGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAATTCGGTGGTGTG 3362
|||||

Qy 61 TCGGGGAAGCTGACGTCCTTTCCATGGCTGCTCGCCTGTGT 101
|||||
Db 3363 TCGGGGAAGCTGACGTCCTTTCCATGGCTGCTCGCCTGTGT 3403
|||||

RESULT 10
ADMA47498
ID ADMA47498 standard; DNA; 8092 BP.
XX
XX ADMA47498;
AC
XX
XX 03-JUN-2004 (first entry)
DT
XX
XX Lysosomal enzyme related DNA #2.
DE
XX
XX Lysosomal enzyme; lysosomal storage disease; Gaucher type I disease;
KW Hurler disease; Sanfilippo disease; gene therapy; ds.
XX
XX Unidentified.
OS
XX US2004023218-A1.
PN
XX
XX 05-FEB-2004.
PD
XX
XX 21-JUN-2002; 2002US-00176066.
PF
XX
XX 21-JUN-2002; 2002US-00176066.
PR
XX
XX (DESM/) DESMARIS N.
PA (HEAR/) HEARD J M.
XX
XX Desmaris N, Heard JM;
PI
XX WPI; 2004-142648/14.
DR
XX
XX New purified nucleic acid molecules capable of expressing a lysosomal
PT enzyme, useful for preventing or treating lysosomal storage diseases
PT (e.g. Gaucher type I disease or Sanfilippo disease) in humans.
XX
XX Disclosure; Page 11-15; 20pp; English.

XX SQ The invention relates to a nucleic acid molecule capable of expressing a
CC lysosomal enzyme. The nucleic acid molecule comprises at least a sequence
CC coding for the lysosomal enzyme and a promoter highly active in the brain
CC inserted upstream from the sequence. Compositions and methods of the
CC invention are useful for preventing or treating lysosomal storage
CC diseases e.g. Gaucher type I disease, Hurler disease (MPSI) or Sanfilippo
CC disease (MPSIII) in humans. The invention is also useful in gene therapy.
XX The present sequence is a DNA related to the invention.

XX SQ Sequence 8092 BP; 1650 A; 2268 C; 2226 G; 1948 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 12; Length 8092;
Best Local Similarity 100.0%; Pred. No. 7.9e-22;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAATTCGGTGGTGTG 60
3747 TGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAATTCGGTGGTGTG 3806
|||||

Qy 61 TCGGGGAAGCTGACGTCCTTTCCATGGCTGCTCGCCTGTGT 101
|||||

Db 3807 TCGGGGAAGCTGACGTCCTTTCCATGGCTGCTCGCCTGTGT 3847

RESULT 11
AAA59091
ID AAA59091 standard; DNA; 8484 BP.
XX
XX AAA59091;
AC
XX
XX 07-NOV-2000 (first entry)
DT
XX
XX Nucleotide sequence of plasmid pDV90.
DE
XX
XX Adenovirus; tripartite leader; adenovirus vector particle; gene delivery;
KW ss.
XX
XX Synthetic.
OS
XX WO200042208-A1.
PN
XX
XX 20-JUL-2000.
PD
XX
XX 14-JAN-2000; 2000WO-EP000265.
PF
XX
XX 14-JAN-1999; 99US-0115920P.
PR
XX
XX (NOVS) NOVARTIS AG.
PA (NOVS) NOVARTIS-ERFINDUNGEN VERW GES MBH.
PA (SCRI) SCRIPPS RES INST.
XX
XX Nemerow GR, Von Seggern DJ, Hallenbeck PL, Stevenson SC;
PI Skripchenko Y;
PI
XX WPI; 2000-476068/41.
DR
XX
XX New nucleic acid comprising an adenovirus tripartite leader nucleotide
PT for producing high-capacity and targeted vectors for adenovirus-based
PT gene therapy.

XX SQ Claim 10; Page 200-204; 212pp; English.

XX The specification describes a nucleic acid molecule comprising an
CC adenovirus (AV) tripartite leader (TPL) nucleotide with a sequence
CC comprising two different TPL exons or three same or different TPL exons.
CC The nucleic acid is used to produce an adenovirus vector particle,
CC deliver an exogenous gene to a target cell, pseudotype recombinant viral
CC vectors, target an adenovirus vector to a cell, produce a modified
CC adenovirus, deliver a heterologous gene to an animal and produce a
CC gutless adenoviral vector particle. The present sequence represents
CC plasmid pDV90, which contains a TPL

XX SQ Sequence 8484 BP; 1996 A; 2238 C; 2125 G; 2125 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 3; Length 8484;
Best Local Similarity 100.0%; Pred. No. 8e-22;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAATTCGGTGGTGTG 60
4192 TGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTGGGCACTGACAATTCGGTGGTGTG 4251
|||||

Qy 61 TCGGGGAAGCTGACGTCCTTTCCATGGCTGCTCGCCTGTGT 101
|||||

Db 4252 TCGGGGAAGCTGACGTCCTTTCCATGGCTGCTCGCCTGTGT 4292
|||||

RESULT 12
ADF48775
ID ADF48775 standard; DNA; 8484 BP.
XX
XX ADF48775;
AC
XX
XX 12-FEB-2004 (first entry)
DT

Db 6513 TGCCTTCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAAATTCGGTGGTGTG 6572

Qy 61 TCGGGGAAGCTGACGCTCTTTCCATGCGTCTGCTCGCTGTG 101
 |||||
 Db 6573 TCGGGGAAGCTGACGCTCTTTCCATGCGTCTGCTCGCTGTG 6613
 |||||

RESULT 14

ADM82791/c

ID ADM82791 standard; cDNA; 9731 BP.

XX AC ADM82791;

XX DT 03-JUN-2004 (first entry)

XX DE DNA repair pathway related retroviral vector cDNA with CMV promoter.

XX DE inducer; inhibitor; DNA repair pathway; anti-HIV; cytostatic; virucide;

XX KW antidiabetic; neuroprotective; retroviral infection; AIDS; HIV infection;

XX KW cancer; human adult T-cell leukaemia; lymphoma;

XX KW feline immunodeficiency virus; Type I diabetes; multiple sclerosis;

XX KW gene therapy; human; cyclic; circular; CMV promoter; ss.

XX OS Unidentified.

XX PN WO2003089573-A2.

XX PD 30-OCT-2003.

XX PF 04-APR-2003; 2003WO-US010302.

XX PR 05-APR-2002; 2002US-0370376P.

XX PA (FISH/) FISHEL R A.

XX PA (YODE/) YODER K E.

XX PI Fishel RA, Yoder KE;

XX DR WPI; 2003-854096/79.

XX DR Screening for compounds that modulate a DNA repair pathway and/or

XX PT retroviral integration, useful for treating retroviral infection,

XX PT comprises determining the amount of a retroviral cDNA circularization in

XX PT the presence of the test compound.

XX PS Claim 73; SEQ ID NO 5; 89pp; English.

XX PS The invention relates to a novel method for screening for inducers or

XX CC inhibitors of a DNA repair pathway by contacting at least one component

XX CC of a DNA repair pathway with a non-circularized retroviral cDNA in the

XX CC presence and absence of a test compound, and determining whether

XX CC circularization of the cDNA is increased or decreased in the presence of

XX CC the test compound. The DNA repair pathway components have the following

XX CC activities: anti-HIV, cytostatic, virucide, antidiabetic, and

XX CC neuroprotective. The method is useful for identifying compounds that

XX CC modulate a DNA repair pathway and/or retroviral activity. The compound is

XX CC used in manufacturing a pharmaceutical composition for the treatment of a

XX CC retroviral infection (e.g. AIDS, HIV infection, cancer, human adult T-

XX CC cell leukaemia, lymphoma, feline immunodeficiency virus, Type I diabetes

XX CC or multiple sclerosis) or for increasing the efficiency of gene delivery

XX CC in a gene therapy. This polynucleotide represents a retroviral cDNA

XX CC sequence of the invention.

XX SQ Sequence 9731 BP; 2444 A; 2412 C; 2548 G; 2327 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 11; Length 9731;

Best Local Similarity 100.0%; Pred. No. 8.2e-22;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGCCTTCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAAATTCGGTGGTGTG 60
 |||||

Db 7288 TGCCTTCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAAATTCGGTGGTGTG 7229
 |||||

Qy 61 TCGGGGAAGCTGACGCTCTTTCCATGCGTCTGCTCGCTGTG 101
 |||||

Db 7228 TCGGGGAAGCTGACGCTCTTTCCATGCGTCTGCTCGCTGTG 7188
 |||||

RESULT 15

ADM82792/c

ID ADM82792 standard; cDNA; 9782 BP.

XX AC ADM82792;

XX DT 03-JUN-2004 (first entry)

XX DE DNA repair pathway related retroviral vector cDNA with MSH2 promoter.

XX DE inducer; inhibitor; DNA repair pathway; anti-HIV; cytostatic; virucide;

XX KW antidiabetic; neuroprotective; retroviral infection; AIDS; HIV infection;

XX KW cancer; human adult T-cell leukaemia; lymphoma;

XX KW feline immunodeficiency virus; Type I diabetes; multiple sclerosis;

XX KW gene therapy; human; cyclic; circular; MSH2 promoter; ss.

XX OS Unidentified.

XX PN WO2003089573-A2.

XX PD 30-OCT-2003.

XX PF 04-APR-2003; 2003WO-US010302.

XX PR 05-APR-2002; 2002US-0370376P.

XX PA (FISH/) FISHEL R A.

XX PA (YODE/) YODER K E.

XX PI Fishel RA, Yoder KE;

XX DR WPI; 2003-854096/79.

XX DR Screening for compounds that modulate a DNA repair pathway and/or

XX PT retroviral integration, useful for treating retroviral infection,

XX PT comprises determining the amount of a retroviral cDNA circularization in

XX PT the presence of the test compound.

XX PS Claim 73; SEQ ID NO 6; 89pp; English.

XX PS The invention relates to a novel method for screening for inducers or

XX CC inhibitors of a DNA repair pathway by contacting at least one component

XX CC of a DNA repair pathway with a non-circularized retroviral cDNA in the

XX CC presence and absence of a test compound, and determining whether

XX CC circularization of the cDNA is increased or decreased in the presence of

XX CC the test compound. The DNA repair pathway components have the following

XX CC activities: anti-HIV, cytostatic, virucide, antidiabetic, and

XX CC neuroprotective. The method is useful for identifying compounds that

XX CC modulate a DNA repair pathway and/or retroviral activity. The compound is

XX CC used in manufacturing a pharmaceutical composition for the treatment of a

XX CC retroviral infection (e.g. AIDS, HIV infection, cancer, human adult T-

XX CC cell leukaemia, lymphoma, feline immunodeficiency virus, Type I diabetes

XX CC or multiple sclerosis) or for increasing the efficiency of gene delivery

XX CC in a gene therapy. This polynucleotide represents a retroviral cDNA

XX CC sequence of the invention.

XX SQ Sequence 9782 BP; 2628 A; 2305 C; 2351 G; 2498 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 11; Length 9782;

Best Local Similarity 100.0%; Pred. No. 8.2e-22;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGCCTTCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAAATTCGGTGGTGTG 60
 |||||

Db 1340 TGCCTTCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAAATTCGGTGGTGTG 1281
 |||||

Qy 61 TCGGGGAAGCTGACGCTCTTTCCATGCGTCTGCTCGCTGTG 101
 |||||

Db 1280 TCGGGGAGCTGACGTCCTTTCCATGGCTGCTCGCCTGTGT 1240

Search completed: July 14, 2005, 07:01:51
Job time : 145.448 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:39:07 ; Search time 756.618 Seconds
(without alignments)
6468.225 Million cell updates/sec

Title: US-09-482-682-65_COPY_8384_8484
Perfect score: 101
Sequence: 1 aggtgtattgtctcatgagc.....gaaagtgcacctgacgtc 101

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_by.*
13: gb_un.*
14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	101	100.0	142	6 AR356490	AR356490 Sequence
C 2	101	100.0	142	6 AR538046	AR538046 Sequence
C 3	101	100.0	228	6 E00019	E00019 DNA coding
C 4	101	100.0	240	1 PMOENDO	M10199 Plasmid pMM
C 5	101	100.0	251	6 E00018	E00018 DNA coding
C 6	101	100.0	251	6 I01644	I01644 Sequence 1
C 7	101	100.0	344	11 HUMUT5345	L18624 Human chrom
C 8	101	100.0	400	6 BD195256	BD195256 Nucleotid
C 9	101	100.0	456	6 E00892	E00892 Synthetic D
C 10	101	100.0	456	6 E01156	E01156 DNA fragmen
C 11	101	100.0	456	6 E01274	E01274 DNA encodin
C 12	101	100.0	456	6 E01302	E01302 DNA encodin
C 13	101	100.0	466	6 AX260098	AX260098 Sequence
C 14	101	100.0	573	6 AX260150	AX260150 Sequence
C 15	101	100.0	693	6 A43586	A43586 Sequence 11
C 16	101	100.0	693	6 ARI16755	ARI16755 Sequence
C 17	101	100.0	998	1 AY559171	AY559171 Pseudomon
C 18	101	100.0	1011	1 SMTMAQGE	X97254 S.marcescen
C 19	101	100.0	1012	2 CEC11F10	Z92776 Caenorhabdi

20	101	100.0	1014	4 CFAJ4121	AR224121 Canis fam
C 21	101	100.0	1027	1 AY589493	AY589493 Escherich
C 22	101	100.0	1040	1 AY538698	AY538698 Serratia
C 23	101	100.0	1040	1 AY538700	AY538700 Serratia
C 24	101	100.0	1040	1 AY538701	AY538701 Serratia
C 25	101	100.0	1040	1 AY538702	AY538702 Serratia
C 26	101	100.0	1041	1 AY538699	AY538699 Serratia
C 27	101	100.0	1042	1 AY394610	AY394610 Klebsiell
C 28	101	100.0	1042	1 ECO308558	ECO308558 Escherich
C 29	101	100.0	1044	1 AY392531	AY392531 Streptoco
C 30	101	100.0	1044	1 AY452662	AY452662 Streptoco
C 31	101	100.0	1054	1 AF104441	AF104441 Klebsiell
C 32	101	100.0	1054	1 AF104442	AF104442 Escherich
C 33	101	100.0	1058	6 I03356	I03356 Sequence 4
C 34	101	100.0	1064	1 AY628199	AY628199 Escherich
C 35	101	100.0	1069	1 AF535127	AF535127 Klebsiell
C 36	101	100.0	1069	1 AY243512	AY243512 Klebsiell
C 37	101	100.0	1071	1 AY628175	AY628175 Escherich
C 38	101	100.0	1072	1 AY101764	AY101764 Klebsiell
C 39	101	100.0	1073	6 AR371489	AR371489 Sequence
C 40	101	100.0	1073	6 AX195443	AX195443 Sequence
C 41	101	100.0	1075	1 AY729027	AY729027 Proteus m
C 42	101	100.0	1075	1 PATN1PN2	X54605 Pseudomonas
C 43	101	100.0	1075	1 PATN2PN1B	X54607 Pseudomonas
C 44	101	100.0	1075	1 PATN3PN1A	X54604 Pseudomonas
C 45	101	100.0	1080	1 AF027199	AF027199 Klebsiell

ALIGNMENTS

RESULT 1
AR356490/c
LOCUS AR356490 142 bp DNA linear PAT 17-AUG-2003
DEFINITION Sequence 2608 from patent US 6593114.
ACCESSION AR356490
VERSION AR356490.1 GI:33762574
KEYWORDS SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 142)
AUTHORS Kunsch,C.A., Choi,G.H., Barash,S., Dillon,P.J., Fannon,M.R. and Rosen,C.A.
TITLE Staphylococcus aureus polynucleotides and sequences
JOURNAL Patent: US 6593114-A 2608 15-JUL-2003;
FEATURES Location/Qualifiers
source
1..142
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 142;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AGGTTATTGCTCATGCGGATACATATTGTAATGATTATTAGAAAATAAACAATAG 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 107 AGGTTATTGCTCATGCGGATACATATTGTAATGATTATTAGAAAATAAACAATAG 48
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Oy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 101
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 47 GGGTTCGCGCACATTTCCCGAAAAGTGCACCTGACGTC 7
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 2
AR538046/c
LOCUS AR538046 142 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 2608 from patent US 6737248.
ACCESSION AR538046
VERSION AR538046.1 GI:53929263
KEYWORDS SOURCE
Unknown.

```

ORGANISM Unknown.
REFERENCE 1 (bases 1 to 142)
AUTHORS Kunsch,C.A., Choi,G.A., Barash,S.C., Dillon,P.J., Fannon,M.R. and Rosen,C.A.
TITLE Staphylococcus aureus polynucleotides and sequences
JOURNAL Patent: US 6737248-A 2608 18-MAY-2004;
FEATURES
    source
        1..142
        /organism="unknown"
        /mol_type="genomic DNA"
ORIGIN
    Query Match 100.0%; Score 101; DB 6; Length 142;
    Best Local Similarity 100.0%; Pred. No. 8.7e-20;
    Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 60
Db 107 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 48
Qy 61 GGGTTCGGCGCACATTTCCCGGAAAAAGTGCCACCTGACGTC 101
Db 47 GGGTTCGGCGCACATTTCCCGGAAAAAGTGCCACCTGACGTC 7
RESULT 3
E00019/c
LOCUS DNA coding for Escherichia coli penicillinase.
DEFINITION E00019
VERSION E00019.1 GI:2168327
KEYWORDS JP 1981154999-A/2.
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
1 (bases 1 to 228)
Uorutaa,G. and Karen,T.
SYNTHESIS OF SELECTIVE AGED PROTEIN OR POLYPEPTIDE IN BACTERIA
PATENT: JP 1981154999-A 2 30-NOV-1981;
JOURNAL UNIV HARVARD
COMMENT OS Escherichia coli
PN JP 1981154999-A/2
PD 30-NOV-1981
PF 09-APR-1981 JP 1981052488
PR 11-APR-1980 US 80 139225
PI UORUTAA GIRUBAATO, KAREN TARUMATSUJI
PC C12P21/00,C07H21/00,C12N1/00,C12N15/00//C12R1/19; CC
strandedness: Double;
CC topology: Linear;
CC anti-sense: No;
CC *source: clone=pKT218;
FH Key Location/Qualifiers
FT CDS 210..>228
    /product='E.coli penicillinase'.
FEATURES
    source
        1..228
        Location/Qualifiers
ORIGIN
    Query Match 100.0%; Score 101; DB 6; Length 228;
    Best Local Similarity 100.0%; Pred. No. 8.7e-20;
    Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 60
Db 175 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 116
Qy 61 GGGTTCGGCGCACATTTCCCGGAAAAAGTGCCACCTGACGTC 101

```

```

Db 115 GGGTTCGGCGCACATTTCCCGGAAAAAGTGCCACCTGACGTC 75
RESULT 4
PMOENDO/c
LOCUS DNA 240 bp linear BCT 26-APR-1993
DEFINITION Plasmid pMM110 region of endo VII cleavage sites near cruciform structures.
ACCESSION M10199
VERSION M10199.1 GI:150826
KEYWORDS
SOURCE Plasmid pMM110
ORGANISM Plasmid pMM110
other sequences, plasmids.
REFERENCE 1 (bases 1 to 240)
AUTHORS Kemper,B., Jensch,F., von Depka-Prondzynski,M., Fritz,H.J., Borgmeyer,U. and Mizuuchi,K.
TITLE Resolution of Holliday structures by endonuclease VII as observed in interactions with cruciform DNA
JOURNAL Cold Spring Harb. Symp. Quant. Biol. 49, 815-825 (1984)
MEDLINE 85153063
PUBMED 6397324
COMMENT Original source text: Plasmid pMM110 DNA.
FEATURES
    source
        1..240
        Location/Qualifiers
        /organism="Plasmid pMM110"
        /mol_type="genomic DNA"
        /db_xref="taxon:2599"
        /plasmid="Plasmid pMM110"
ORIGIN Unreported.
    Query Match 100.0%; Score 101; DB 1; Length 240;
    Best Local Similarity 100.0%; Pred. No. 8.7e-20;
    Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 60
Db 151 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 92
Qy 61 GGGTTCGGCGCACATTTCCCGGAAAAAGTGCCACCTGACGTC 101
Db 91 GGGTTCGGCGCACATTTCCCGGAAAAAGTGCCACCTGACGTC 51
RESULT 5
E00018/c
LOCUS DNA coding for Escherichia coli penicillinase.
DEFINITION E00018
ACCESSION E00018.1 GI:2168326
VERSION JP 1981154999-A/1.
KEYWORDS Escherichia coli
SOURCE Escherichia coli
ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
1 (bases 1 to 251)
Uorutaa,G. and Karen,T.
SYNTHESIS OF SELECTIVE AGED PROTEIN OR POLYPEPTIDE IN BACTERIA
PATENT: JP 1981154999-A 1 30-NOV-1981;
JOURNAL UNIV HARVARD
COMMENT OS Escherichia coli
PN JP 1981154999-A/1
PD 30-NOV-1981
PF 09-APR-1981 JP 1981052488
PR 11-APR-1980 US 80 139225
PI UORUTAA GIRUBAATO, KAREN TARUMATSUJI
PC C12P21/00,C07H21/00,C12N1/00,C12N15/00//C12R1/19; CC
strandedness: Double;
CC topology: Linear;
CC anti-sense: No;
CC fragment type: N-Terminal Fragment;
CC *source: Clone=pKT241;

```

FH Key Location/Qualifiers
FH CDS 210..>252
FT /product='E.coli penicillinase' FT
TATA_signal 190..196
FEATURES
source Location/Qualifiers
1..251
/organism='Escherichia coli'
/mol_type='genomic DNA'
/db_xref='taxon:562'
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 251;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAAATAG 60
|||||
Db 175 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAAATAG 116
|||||
QY 61 GGGTTCGCGCACATTTCCCGGAAAGTGCACCTGACGTC 101
|||||
Db 115 GGGTTCGCGCACATTTCCCGGAAAGTGCACCTGACGTC 75
|||||
RESULT 6
I01644/c
LOCUS 251 bp ss-DNA linear PAT 18-MAY-1993
DEFINITION Sequence 1 from Patent US 4338397.
ACCESSION I01644
VERSION I01644.1 GI:267685
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 251)
AUTHORS Gilbert, W. and Talmadge, K.
TITLE Mature protein synthesis
JOURNAL Patent: US 4338397-A 1 06-JUL-1982;
President and Fellows of Harvard College; Cambridge, MA
FEATURES
source Location/Qualifiers
1..251
/organism='unknown'
/mol_type='unassigned DNA'
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 251;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAAATAG 60
|||||
Db 175 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAAATAG 116
|||||
QY 61 GGGTTCGCGCACATTTCCCGGAAAGTGCACCTGACGTC 101
|||||
Db 115 GGGTTCGCGCACATTTCCCGGAAAGTGCACCTGACGTC 75
|||||
RESULT 7
HUMUT5345
LOCUS 344 bp DNA linear STS 26-JUL-1993
DEFINITION Human chromosome 8 STS UT5345, sequence tagged site.
ACCESSION L18624
VERSION L18624.1 GI:308338
KEYWORDS STS; PCR primer; STS sequence; microsatellite marker;
microsatellite repeat; repeat polymorphism; sequence tagged site.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 344)
AUTHORS Gerken, S.C., Matsunami, N., Lawrence, E., Carlson, M., Moore, M.,

Ballard, L., Melis, R., Robertson, M., Bradley, P., Elsner, T.,
Tingey, A., Rodriguez, P., Albertsen, H., Lalouel, J.-M. and White, R.
Genetic and physical mapping of simple sequence repeat containing
sequence tagged sites from the human genome
Unpublished (1993)
Original source text: Homo sapiens DNA.
Submitted by: Utah Center for Human Genome Research University of
Utah, Dept. of Human Genetics
2160 Eccles Institute of Human Genetics
Salt Lake City, UT 84112
e-mail: sts@corona.med.utah.edu
Primer A: GAGCAAAACAGGAGGCAAAATGC
Primer B: TTCGGGAATGTGCGCGAACC
32P-label: B Primer
PCR Profile:
Initial Denaturation: 94C 300sec
PCR Cycles: 30
Denaturation: 94C 10sec
Annealing: 60C 10sec
Extension: 72C 20sec
Mg++: 2mM
Gel: Acrylamide 7%, Formamide 32%, Urea 34%
Alleles: 2
FEATURES
source Location/Qualifiers
1..344
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'
/map='8'
36..224
/standard_name='STS UT5345'
36..60
complement(202..224)
ORIGIN
Query Match 100.0%; Score 101; DB 11; Length 344;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAAATAG 60
|||||
Db 141 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAAATAG 200
|||||
QY 61 GGGTTCGCGCACATTTCCCGGAAAGTGCACCTGACGTC 101
|||||
Db 201 GGGTTCGCGCACATTTCCCGGAAAGTGCACCTGACGTC 241
|||||
RESULT 8
BD195256/c
LOCUS 400 bp DNA linear PAT 17-JUL-2003
DEFINITION Nucleotide sequence of Escherichia coli pathogenicity islands.
ACCESSION BD195256
VERSION BD195256.1 GI:33005021
KEYWORDS JP 2002513277-A/43.
SOURCE unidentified
ORGANISM unidentified
unclassified.
REFERENCE 1 (bases 1 to 400)
AUTHORS Dillon, P.J., Choi, G.H. and Welch, R.A.
TITLE Nucleotide sequence of Escherichia coli pathogenicity islands
JOURNAL Patent: JP 2002513277-A 43 08-MAY-2002;
HUMAN GENOME SCIENCES INC, WISCONSIN ALUMNI RESEARCH FOUNDATION
COMMENT OS Unidentified
PN JP 2002513277-A/43
PD 08-MAY-2002
PF 21-NOV-1997 JP 1998523916
PR 60/031626, 14-OCT-1997 US 60/061953 PI
PATRICK J DILLON, GIL H CHOI, RODNEY A WELCH
PC C12N15/11, C12N15/63, C07K16/12, G01N33/569, G06F17/30, G11B7/00 CC
Strandedness: Double;
CC Topology: Linear;
CC Nucleotide sequence of Escherichia coli pathogenicity islands

```

FH Key Location/Qualifiers
FT source 1..400
FT /organism='Unidentified'.
FEATURES
  source 1..400
    Location/Qualifiers
    1..400
    /organism='unidentified'
    /mol_type='genomic DNA'
    /db_xref='taxon:32644'
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 400;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGTTATTGTCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 60
Db 165 AGGTTATTGTCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 106
Qy 61 GGGTTCGCGCACATTTCCCGGAAAAGTGCACCTGACGTC 101
Db 105 GGGTTCGCGCACATTTCCCGGAAAAGTGCACCTGACGTC 65
RESULT 9
E00892/c
LOCUS E00892 Synthetic DNA encoding fused polypeptide between E coli 29-SEP-1997
DEFINITION beta-lactamase and human beta-urogastrone.
ACCESSION E00892
VERSION E00892.1 GI:2169153
KEYWORDS JP 1986149089-A/1.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 456)
AUTHORS Okai,H., Momota,Y., Kumakura,T., Tochifusa,N., Kitazawa,T.,
TITLE Ojida,K. and Matsushiro,A.
POLYPEPTIDE SECRETION DEVELOPMENT VECTOR, MICROORGANISM TRANSFORMED
WITH SAME, AND PRODUCTION OF POLYPEPTIDE WITH MICROORGANISM
PATENT: JP 1986149089-A 1 07-JUL-1986;
JOURNAL EARTH CHEM CORP LTD
COMMENT OS Artificial gene
OC Artificial sequence; Genes.
PN JP 1986149089-A/1
PD 07-JUL-1986
PF 21-DEC-1984 JP 1984271206
PI OKAI HIDEO, MOMOTA YUTAKA, KUMAKURA TAKESHI,
PI TOCHIFUSA NORIYUKI,
PI KITAZAWA TOSHIKI, OJIDA KAZUHIRO, MATSUSHIRO AIZO
PC C12N15/00,C12N1/20,C12P21/00,(C12N1/20,C12R1:19),(C12P21/00,PC
C12R1:19);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: strain=HB101;
CC *source: clone=pVG201;
CC Feature is identified by experimental;
FH Key Location/Qualifiers
FT promoter 125..170
FT of beta-lactamase
FT RBS 200..203
FT CDS 209..438
FT sig_peptide 209..277
FT mat_peptide 278..435
FT /product='beta-urogastrone precursor' FT
FT /product='signal peptide of beta-lactonase' FT
FT /product='beta-urogastrone mature peptide'.
FEATURES
  source 1..456
    Location/Qualifiers
    1..456
    /organism='synthetic construct'
    /mol_type='genomic DNA'

```

```

/db_xref='taxon:32630'
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGTTATTGTCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 60
Db 173 AGGTTATTGTCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 114
Qy 61 GGGTTCGCGCACATTTCCCGGAAAAGTGCACCTGACGTC 101
Db 113 GGGTTCGCGCACATTTCCCGGAAAAGTGCACCTGACGTC 73
RESULT 10
E01156/c
LOCUS E01156 DNA fragment which secrets beta urogastrone. 29-SEP-1997
DEFINITION E01156
ACCESSION E01156.1 GI:2169415
VERSION JP 1987083890-A/1.
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 456)
AUTHORS Yoshikawa,K., Momota,Y., Kajifusa,N., Koide,T. and Okai,H.
TITLE POLYPEPTIDE SECRETION EXPRESSION VECTOR, MICROORGANISM TRANSFORMED
BY SAID VECTOR AND PRODUCTION OF POLYPEPTIDE USING SAID
PATENT: JP 1987083890-A 1 17-APR-1987;
JOURNAL EARTH CHEM CORP LTD
COMMENT OS Artificial gene
OC Artificial sequence; Genes.
PN JP 1987083890-A/1
PD 17-APR-1987
PF 09-OCT-1985 JP 1985225393
PI YOSHIKAWA KAZUTOSHI, MOMOTA YUTAKA, KAJIFUSA NORIYUKI, PI
KOIDE TAKAO,
PI OKAI HIDEO
PC C12N15/00,C12N1/20,C12P21/00,(C12N1/20,C12R1:19),(C12N1/20,PC
C12R1:125);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: clone=pUG201;
FH Key Location/Qualifiers
FT promoter 125..170
FT /note='beta lactamase promoter' FT RBS
FT CDS 209..439
FT /product='beta urogastrone'
FT sig_peptide 209..277
FT mat_peptide 278..436
FT /product='beta urogastrone'.
FEATURES
  source 1..456
    Location/Qualifiers
    1..456
    /organism='synthetic construct'
    /mol_type='genomic DNA'
    /db_xref='taxon:32630'
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGTTATTGTCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 60
Db 173 AGGTTATTGTCTCATGAGCGGATACATATTGTAATGTTATTAGAAAAATAAACAAATAG 114
Qy 61 GGGTTCGCGCACATTTCCCGGAAAAGTGCACCTGACGTC 101

```

```
Db 113 GGGTTCGCGCACATTTCCTCCCGAAAGTGCACCTGACGTC 73
|||||
RESULT 11
E01274/c
LOCUS E01274 456 bp DNA linear PAT 29-SEP-1997
DEFINITION DNA encoding beta-urogastron fused with DNA encoding a promoter and
signal peptide of beta-lactamase.
ACCESSION E01274
VERSION E01274.1 GI:2169533
KEYWORDS JP 1987179398-A/1.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 456)
AUTHORS Okai,H., Kumakura,T., Kawamoto,S., Adachi,S., Matsuura,A.,
Ojida,K., Yano,M., Mihara,S., Matsushiro,A. and Yanaiharu,N.
TITLE PRODUCTION OF BETA-UROGASTRONE
JOURNAL Patent: JP 1987179398-A 1 06-AUG-1987;
COMMENT EARTH CHEM CORP LTD
OS Artificial gene
OC Artificial sequence; Genes.
OS Homo sapiens
PN JP 1987179398-A/1
PD 06-AUG-1987
PF 31-JAN-1986 JP 1986021032
PI OKAI HIDEO, KUMAKURA TAKESHI, KAWAMOTO SHOICHI, ADACHI SHOJI,
MATSUBARA AKIMASA, OJIDA KAZUHIRO, YANO MAKI, MIHARA SHIGERU,
MATSUSHIRO AIZO, YANAIHARA NOBORU
PC C12P21/00,C12N15/00,(C12P21/00,C12R1:91);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No; Location/Qualifiers
FH Key
FT Promoter 125..170
FT RBS 200..203
FT sig_peptide 209..277
FT mat_peptide 278..436
FT CDS 209..439
FT /product='beta-urogastron'
FEATURES
source Location/Qualifiers
1..456
/organism='synthetic construct'
/mol_type='genomic DNA'
/db_xref='taxon:32630'
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGAATGATTATAGAAAATAAACAATAG 60
|||||
Db 173 AGGGTTATTGCTCATGAGCGGATACATATTGAATGATTATAGAAAATAAACAATAG 114
|||||
Qy 61 GGGTTCGCGCACATTTCCTCCCGAAAGTGCACCTGACGTC 101
|||||
Db 113 GGGTTCGCGCACATTTCCTCCCGAAAGTGCACCTGACGTC 73
|||||
RESULT 12
E01302/c
LOCUS E01302 456 bp DNA linear PAT 29-SEP-1997
DEFINITION DNA encoding human beta-urogastrone fused with DNA encoding
promoter and signal peptide of beta-lactamase.
ACCESSION E01302
VERSION E01302.1 GI:2169561
KEYWORDS JP 1987190083-A/1.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 456)
AUTHORS Deak,P., Glover,D.M. and Midgley,C.
TITLE Cell cycle progression proteins
JOURNAL Patent: WO 0172774-A 60 04-OCT-2001;
Cyclacel Limited (GB)
FEATURES
source Location/Qualifiers
1..466
/organism='Drosophila melanogaster'
/mol_type='unassigned DNA'
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGAATGATTATAGAAAATAAACAATAG 60
|||||
Db 173 AGGGTTATTGCTCATGAGCGGATACATATTGAATGATTATAGAAAATAAACAATAG 114
|||||
Qy 61 GGGTTCGCGCACATTTCCTCCCGAAAGTGCACCTGACGTC 101
|||||
Db 113 GGGTTCGCGCACATTTCCTCCCGAAAGTGCACCTGACGTC 73
|||||
RESULT 13
AX260098/c
LOCUS AX260098 466 bp DNA linear PAT 26-OCT-2001
DEFINITION Sequence 60 from Patent WO0172774.
ACCESSION AX260098
VERSION AX260098.1 GI:16509129
KEYWORDS Drosophila melanogaster (fruit fly)
SOURCE Drosophila melanogaster
ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1
AUTHORS Deak,P., Glover,D.M. and Midgley,C.
TITLE Cell cycle progression proteins
JOURNAL Patent: WO 0172774-A 60 04-OCT-2001;
Cyclacel Limited (GB)
FEATURES
source Location/Qualifiers
1..466
/organism='Drosophila melanogaster'
/mol_type='unassigned DNA'
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGAATGATTATAGAAAATAAACAATAG 60
|||||
Db 173 AGGGTTATTGCTCATGAGCGGATACATATTGAATGATTATAGAAAATAAACAATAG 114
|||||
Qy 61 GGGTTCGCGCACATTTCCTCCCGAAAGTGCACCTGACGTC 101
|||||
Db 113 GGGTTCGCGCACATTTCCTCCCGAAAGTGCACCTGACGTC 73
|||||
```

```
/db_xref="taxon:7227"

ORIGIN
Query Match      100.0%; Score 101; DB 6; Length 466;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 60
    |||||||
Db 280 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 221
    |||||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 101
    |||||||
Db 220 GGGTTCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 180
    |||||||

RESULT 14
AX260150/c      AX260150      573 bp      DNA      linear      PAT 26-OCT-2001
LOCUS           Sequence 112 from Patent WO0172774.
ACCESSION       AX260150
VERSION         AX260150.1 GI:16509172
KEYWORDS
SOURCE
ORGANISM        Drosophila melanogaster (fruit fly)
                Drosophila melanogaster
                Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
                Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
                Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
AUTHORS         Deak, P., Glover, D.M. and Midgley, C.
TITLE           Cell cycle progression proteins
JOURNAL         Patent: WO 0172774-A 112 04-OCT-2001;
                Cyclacel Limited (GB)
FEATURES
  source
    1..573
    /organism="Drosophila melanogaster"
    /mol_type="unassigned DNA"
    /db_xref="taxon:7227"

ORIGIN
Query Match      100.0%; Score 101; DB 6; Length 573;
Best Local Similarity 100.0%; Pred. No. 8.8e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 60
    |||||||
Db 355 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 296
    |||||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 101
    |||||||
Db 295 GGGTTCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 255
    |||||||

RESULT 15
A43586
LOCUS           A43586      693 bp      DNA      linear      PAT 06-MAR-1997
DEFINITION      Sequence 11 from Patent WO9507357.
ACCESSION       A43586
VERSION         A43586.1 GI:2298779
KEYWORDS
SOURCE
ORGANISM        Cuphea lanceolata
                Cuphea lanceolata
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                rosids; Myrtales; Lythraceae; Cuphea.
                1 (bases 1 to 693)
REFERENCE
AUTHORS         Toepfer, R., Bautor, J., Bothmann, H., Filsak, E.,
                Hoerlcke-Grandpierre, C., Klein, B., Martini, N., Mueller, A.,
                Schulte, W., Voetz, M., Walek, J. and Schell, J.
                PROMOTERS
TITLE           Patent: WO 9507357-A 11 16-MAR-1995;
                MAX PLANCK GESELLSCHAFT (DE)
JOURNAL         Other publication CA 2169093 950316
COMMENT
```

```
FEATURES
  source
    Other publication AU 7615494 950327.
    Location/Qualifiers
      1..693
      /organism="Cuphea lanceolata"
      /mol_type="unassigned DNA"
      /db_xref="taxon:3930"
      /clone="CLKASIG8"
      /clone_lib="Genomic Lambda Fix II"

ORIGIN
Query Match      100.0%; Score 101; DB 6; Length 693;
Best Local Similarity 100.0%; Pred. No. 8.8e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 60
    |||||||
Db 592 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 651
    |||||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 101
    |||||||
Db 652 GGGTTCGCGCACATTTCCCGAAAAAGTGCCACCTGACGTC 692
    |||||||

Search completed: July 14, 2005, 14:03:38
Job time : 757.618 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:35:42 ; Search time 142.398 Seconds
(without alignments)
4198.742 Million cell updates/sec

Title: US-09-482-682-65_COPY_8384_8484
Perfect score: 101
Sequence: 1 aggtttattgtctcatgacg.....gaaaagtccacctgacgtc 101

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:.*
1: geneseqn1980s:.*
2: geneseqn1990s:.*
3: geneseqn2000s:.*
4: geneseqn2001as:.*
5: geneseqn2001bs:.*
6: geneseqn2002as:.*
7: geneseqn2002bs:.*
8: geneseqn2003as:.*
9: geneseqn2003bs:.*
10: geneseqn2003cs:.*
11: geneseqn2003ds:.*
12: geneseqn2004as:.*
13: geneseqn2004bs:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	101	100.0	142	2	Aav76919 Staphyloc
C 2	101	100.0	228	1	Aan10032 Sequence
C 3	101	100.0	251	1	Aan10031 Sequence
C 4	101	100.0	400	2	Aav31229 E. coli J
C 5	101	100.0	456	1	Aan60624 Plasmid p
C 6	101	100.0	456	1	Aan71080 Sequence
C 7	101	100.0	456	1	Aan70833 Beta-urog
C 8	101	100.0	456	1	Aan81765 Sequence
C 9	101	100.0	466	6	Aba90413 Drosophil
C 10	101	100.0	487	2	Aax21173 Polynucle
C 11	101	100.0	535	2	Aax21149 Polynucle
C 12	101	100.0	573	6	Aba90456 Drosophil
C 13	101	100.0	605	12	Adh58311 Electroph
C 14	101	100.0	776	4	Aas30560 DNA encod
C 15	101	100.0	776	4	Aas27819 DNA encod
C 16	101	100.0	776	4	Abk42984 Genomic s
C 17	101	100.0	776	4	Aal07344 Human rep
C 18	101	100.0	776	4	Aal03229 Human rep
C 19	101	100.0	776	4	Aal06588 Human rep
C 20	101	100.0	776	4	Aal07340 Human rep

C 21	101	100.0	776	5	ABA14573
C 22	101	100.0	776	5	AAS34681
C 23	101	100.0	776	8	ADA41574
C 24	101	100.0	776	8	ACC50905
C 25	101	100.0	776	8	ABZ71508
C 26	101	100.0	776	9	ADB91869
C 27	101	100.0	776	9	ADB61140
C 28	101	100.0	776	10	ADB94622
C 29	101	100.0	776	10	ADC74663
C 30	101	100.0	776	10	ADA57709
C 31	101	100.0	776	12	ADN41551
C 32	101	100.0	845	4	AAS30559
C 33	101	100.0	845	4	AAS27818
C 34	101	100.0	845	4	ABK42983
C 35	101	100.0	845	4	AAS41807
C 36	101	100.0	845	4	AAS41855
C 37	101	100.0	845	4	AAK85485
C 38	101	100.0	845	4	AAK85434
C 39	101	100.0	845	4	AAL07343
C 40	101	100.0	845	4	AAL06587
C 41	101	100.0	845	4	AAL07339
C 42	101	100.0	845	4	AAL03228
C 43	101	100.0	845	5	ABA14572
C 44	101	100.0	845	5	AAS34680
C 45	101	100.0	845	9	ADB61139

ALIGNMENTS

RESULT 1
AAV76919/c
ID AAV76919 standard; DNA; 142 BP.
XX
AC AAV76919;
DT 16-MAR-1999 (first entry)
XX
DB Staphylococcus aureus contig SEQ ID #2608.
XX
KW Computer readable medium; vaccine; S.aureus infection; immunodetection;
KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW skin infection; surgical wound infection; scalded skin syndrome;
KW toxic shock syndrome; ds.
XX
OS Staphylococcus aureus.
XX
FN EP786519-A2.
XX
PD 30-JUL-1997.
XX
PF 07-JAN-1997; 97EP-00100117.
XX
PR 05-JAN-1996; 96US-0009861P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Kunsch CA, Choi GH, Barash SC, Dillon PJ, Fannon MR, Rosen CA;
XX
XX WPI; 1997-374922/35.
XX
DR Polynucleotide(s) and proteins derived from Staphylococcus aureus -
XX stored on computer readable medium and used in the production of anti-
XX S.aureus vaccines.
XX
PS Claim 1; Page 2287; 3271pp; English.
XX
CC This sequence represents one of 5191 Staphylococcus aureus DNA sequences
CC of the invention. The DNA sequences are recorded on a computer readable
CC medium, preferably selected from a floppy or hard disk, random access
CC memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
CC the S.aureus DNA sequences allows putative functions to be assigned so
CC that protein-encoding or regulatory regions of commercial, therapeutic or

CC industrial importance can be obtained. Specifically, sequences which are
 CC likely to encode antigens have been identified and these polypeptides can
 CC be used in a vaccine composition against *S. aureus* infection. The
 CC polypeptides can also be used in a kit for the immunodetection of
 CC *S. aureus* in a sample. *S. aureus* is implicated in numerous human diseases,
 CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,
 CC skin and surgical wound infections, scalded skin syndrome, toxic shock
 CC syndrome, etc. Organisms transformed with the DNA sequences can be used
 CC for recombinant production of the polypeptides. The new DNA sequences
 CC (and their fragments) are useful as primers or probes for isolating
 CC homologues of any of the *S. aureus* DNA sequences contained on the computer
 CC readable medium

SQ Sequence 142 BP; 45 A; 25 C; 26 G; 45 T; 0 U; 1 Other;

Query Match 100.0%; Score 101; DB 2; Length 142;

Best Local Similarity 100.0%; Pred. No. 2.1e-21;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGCGGATACATATTTGAATGATTTAGAAAAATAAACAAATAG 60

Db 107 AGGGTTATTGCTCATGCGGATACATATTTGAATGATTTAGAAAAATAAACAAATAG 48

Qy 61 GGGTTCCGCGCACATTTCCCGAAAAAGTGCACCTGACGTC 101

Db 47 GGGTTCCGCGCACATTTCCCGAAAAAGTGCACCTGACGTC 7

RESULT 2

AAN10032/C

ID AAN10032 standard; DNA; 228 BP.

XX AC AAN10032;

XX DT 13-AUG-1992 (first entry)

XX DE Sequence of the pKT218 EcoRI-PstI penicillinase gene fragment.

XX KW Cloning vehicle; bacterial vector; transformed host; penicillinase;

XX KW insulin; ds.

XX OS Escherichia coli.

XX FT Key Location/Qualifiers

XX FT misc_feature 1..4

XX FT /*tag= a

XX FT /label= sticky end

XX FT 225..228

XX FT /*tag= b

XX FT /label= sticky end

XX EP38182-A.

XX 21-OCT-1981.

XX 09-APR-1981; 81EP-00301561.

XX 11-APR-1980; 80US-00139225.

XX (HARD) HARVARD COLLEGE.

XX Gilbert W, Talmadge K;

XX WPI; 1981-80125D/44.

XX P-PSDB; AAP10039.

XX Synthesis of mature protein or polypeptide - by using bacterial host

XX transformed by cloned vehicle contg. DNA fragment etc.

XX Example; Fig 3; 34pp; English.

XX The closest identifiable promoter for the penicillinase gene in pKT241

XX (AAN10031) is located in the region 14 to 20 nucleotides before its

CC translational start signal. In the examples, the 3' end of pKT241 was
 CC attached to the signal DNA sequence of the DNA fragment (19) for rat
 CC preproinsulin (see AAN10033). The closest identifiable promoter for the
 CC penicillinase gene in pKT218 (AAN10032) is located in the region 14 to 20
 CC nucleotides before its translational start signal. In the examples, the
 CC 3' end of pKT218 was attached to the signal DNA sequence of the DNA
 CC fragment (CB6) for rat preproinsulin (see AAN10034)

SQ Sequence 228 BP; 72 A; 37 C; 46 G; 73 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 1; Length 228;

Best Local Similarity 100.0%; Pred. No. 2.3e-21;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGCGGATACATATTTGAATGATTTAGAAAAATAAACAAATAG 60

Db 175 AGGGTTATTGCTCATGCGGATACATATTTGAATGATTTAGAAAAATAAACAAATAG 116

Qy 61 GGGTTCCGCGCACATTTCCCGAAAAAGTGCACCTGACGTC 101

Db 115 GGGTTCCGCGCACATTTCCCGAAAAAGTGCACCTGACGTC 75

RESULT 3

AAN10031/C

ID AAN10031 standard; DNA; 251 BP.

XX AC AAN10031;

XX DT 13-AUG-1992 (first entry)

XX DE Sequence of the pKT241 EcoRI-PstI penicillinase gene fragment.

XX KW Cloning vehicle; bacterial vector; transformed host; penicillinase;

XX KW insulin; ds.

XX OS Escherichia coli.

XX FT Key Location/Qualifiers

XX FT misc_feature 1..4

XX FT /*tag= a

XX FT /label= sticky end

XX FT 248..251

XX FT /*tag= b

XX FT /label= sticky end

XX EP38182-A.

XX 21-OCT-1981.

XX 09-APR-1981; 81EP-00301561.

XX 11-APR-1980; 80US-00139225.

XX (HARD) HARVARD COLLEGE.

XX Gilbert W, Talmadge K;

XX WPI; 1981-80125D/44.

XX P-PSDB; AAP10038.

XX Synthesis of mature protein or polypeptide - by using bacterial host

XX transformed by cloned vehicle contg. DNA fragment etc.

XX Example; Fig 2; 34pp; English.

XX The closest identifiable promoter for the penicillinase gene in pKT241

XX (AAN10031) is located in the region 14 to 20 nucleotides before its

XX translational start signal. In the examples, the 3' end of pKT241 was

XX attached to the signal DNA sequence of the DNA fragment (19) for rat

XX preproinsulin (see AAN10033). The closest identifiable promoter for the

XX penicillinase gene in pKT218 (AAN10032) is located in the region 14 to 20

XX nucleotides before its translational start signal. In the examples, the

CC 3' end of pKT218 was attached to the signal DNA sequence of the DNA
CC fragment (CB6) for rat preproinsulin (see AAN10034)
XX
SQ Sequence 251 BP; 74 A; 45 C; 50 G; 82 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 1; Length 251;
Best Local Similarity 100.0%; Pred. No. 2.3e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGCGCGATACATATTGAATGTTAGAAAAATAAACAATAG 60
Db |||||
175 AGGGTTATTGCTCATGCGCGATACATATTGAATGTTAGAAAAATAAACAATAG 116
|||

Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101
Db |||||
115 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 75
|||

RESULT 4
AAV31229/c
ID AAV31229 standard; DNA; 400 BP.
XX AC AAV31229;
XX
DT 01-OCT-1998 (first entry)
XX
DE E. coli J96 pathogenicity island contig #43.
XX
KW PAI; pathogenicity island; uropathogenic E. coli detection; PAI IV; pHER;
KW PAI V; phev; vaccine; protective immune response; ds.
XX
OS Escherichia coli.
XX
XX WO9822575-A2.
XX
PD 28-MAY-1998.
XX
XX 21-NOV-1997; 97WO-US021347.
XX
PR 22-NOV-1996; 96US-0031626P.
PR 14-OCT-1997; 97US-0061953P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA (UYWI-) UNIV WISCONSIN.
XX
XX Dillon PJ, Choi GH, Welch RA;
XX WPI; 1998-312461/27.
XX
XX New isolated uropathogenic E. coli nucleotide sequences - used to develop
PT products for the detection of pathogenic E. coli and to elicit an immune
PT response to pathogenic E. coli.
XX
PS Claim 21; Page 140-141; 250pp; English.
XX

This sequence represents a E. coli strain J96 contig containing
CC pathogenicity island (PAI) sequences, and represents a nucleic acid
CC molecule of the invention. PAIs are large fragments of DNA which comprise
CC pathogenicity determinants. The sequences of the invention are taken from
CC PAI IV and PAI V. PAI IV is located at approximately 64 min (near phev)
CC on the E. coli chromosome and is greater than 170 kb. PAI V is located at
CC approximately 94 min (at pHER) on the E. coli chromosome and is
CC approximately 160 kb in size. Antibodies specific to the proteins encoded
CC by the PAI open reading frames of the invention can be used in kits to
CC detect uropathogenic E. coli. The proteins are used in vaccines to elicit
CC a protective immune response in an animal to the uropathogenic E. coli
CC strain J96
XX

SQ Sequence 400 BP; 106 A; 77 C; 91 G; 126 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 2; Length 400;
Best Local Similarity 100.0%; Pred. No. 2.5e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGCGCGATACATATTGAATGTTAGAAAAATAAACAATAG 60
Db |||||
165 AGGGTTATTGCTCATGCGCGATACATATTGAATGTTAGAAAAATAAACAATAG 106
|||

Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101
Db |||||
105 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 65
|||

RESULT 5
AAN60624/c
ID AAN60624 standard; DNA; 456 BP.
XX AC AAN60624;
XX
DT 25-MAR-2003 (revised)
DT 29-OCT-1991 (first entry)
XX
DE Plasmid pUG201 sequence encoding beta-urogastrone.
XX
KW Beta-lactamase signal peptide; pGH54; pGH55; ss.
XX
OS Synthetic.
XX

Key Location/Qualifiers
FH 125..170
FT promoter /*tag= a
FT RBS 200..203 /*tag= b
FT CDS 209..439 /*tag= c
FT sig_peptide 209..277 /*tag= d
FT /label= Beta-lactamase signal peptide
FT mat_peptide 278..436 /*tag= e
FT /label= Beta-urogastrone
XX
XX WO8603779-A.
XX
XX 03-JUL-1986.
XX
XX 19-DEC-1985; 85WO-JP000696.
XX
XX 21-DEC-1984; 84JP-00271206.
XX
XX (EART) EARTH CHEM CO LTD.
PA (OHGA/) OHGAI H.
XX
XX Ohgai H, Momota H, Kuwakura T, Kajifusa N, Kitazawa T, Oshiden K;
XX WPI; 1986-182911/28.
DR P-PSDB; AAP60678.
XX

Recombinant vector for polypeptide secretion - contains signal peptide
PT sequence directly bonded to peptide-coding sequence.
XX
PS Disclosure; Table 4; 79pp; Japanese.
XX
CC The plasmid produces secreted beta-urogastrone in a transformed
CC expression system. Similar plasmids may be constructed where the
CC secretion signal may be coupled with eg. somatostatin, insulin, growth
CC hormone, interferon, IL-2, gastric inhibitory peptide, influenza B SA,
CC epidermal growth factor and thymosine-beta4. (Updated on 25-MAR-2003 to
CC correct PA field.)
XX
SQ Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 1; Length 456;
Best Local Similarity 100.0%; Pred. No. 2.6e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTATTTAGAAAAATAAACAATAG 60
Db 173 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTATTTAGAAAAATAAACAATAG 114

Qy 61 GGGTTCGCGCACATTTCCCGGAAAGTGCACCTGACGTC 101
Db 113 GGGTTCGCGCACATTTCCCGGAAAGTGCACCTGACGTC 73

RESULT 6
AAN71080/c
ID AAN71080 standard; DNA; 456 BP.
XX
AC AAN71080;
XX
DT 25-MAR-2003 (revised)
DT 10-MAR-2003 (revised)
DT 13-MAY-1991 (first entry)
XX
DE Sequence encoding beta-urogastrone.
XX
XX pUGT 150s; beta-UG; ds.
XX
KW Escherichia coli.
XX
OS Synthetic.
XX
XX Key Location/Qualifiers
FT promoter 125..170
FT FT /*tag= a
FT CDS 209..439
FT FT /*tag= b
FT FT /*tag= b
FT FT /trans_except= (pos:434..436,aa:Arg)
XX
XX JP62190083-A.
XX
XX 20-AUG-1987.
XX
XX 14-FEB-1986; 86JP-00031415.
XX
XX 14-FEB-1986; 86JP-00031415.
XX
XX (EART ) EARTH SEIYAKU KK.
XX
XX WPI; 1987-273761/39.
XX
XX Expression vector contg. multiple information units is used to transform
XX host all for increased prodn. of polypeptides.
XX
XX Disclosure; Page 553; 34pp; Japanese.
XX
XX Sequence encodes beta-urogastrone under the control of a tac promoter.
XX The peptide may be expressed from plasmid pUGT 150s in a transformed
XX E.coli host. The plasmid may carry several separately expressing
XX sequences comprising a tac promoter, SD site, signal peptide, and coding
XX sequence, to produce beta-UG in high yield. (Updated on 10-MAR-2003 to
XX add missing OS field.) (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 1; Length 456;
Best Local Similarity 100.0%; Pred. No. 2.6e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTATTTAGAAAAATAAACAATAG 60
Db 173 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTATTTAGAAAAATAAACAATAG 114

Qy 61 GGGTTCGCGCACATTTCCCGGAAAGTGCACCTGACGTC 101
Db 113 GGGTTCGCGCACATTTCCCGGAAAGTGCACCTGACGTC 73

RESULT 7
AAN71080/c
ID AAN71080 standard; DNA; 456 BP.
XX
AC AAN71080;
XX
DT 25-MAR-2003 (revised)
DT 10-MAR-2003 (revised)
DT 13-MAY-1991 (first entry)
XX
DE Sequence encoding beta-urogastrone.
XX
XX pUGT 150s; beta-UG; ds.
XX
KW Escherichia coli.
XX
OS Synthetic.
XX
XX Key Location/Qualifiers
FT promoter 125..170
FT FT /*tag= a
FT CDS 209..439
FT FT /*tag= b
FT FT /*tag= b
FT FT /trans_except= (pos:434..436,aa:Arg)
XX
XX JP62190083-A.
XX
XX 20-AUG-1987.
XX
XX 14-FEB-1986; 86JP-00031415.
XX
XX 14-FEB-1986; 86JP-00031415.
XX
XX (EART ) EARTH SEIYAKU KK.
XX
XX WPI; 1987-273761/39.
XX
XX Expression vector contg. multiple information units is used to transform
XX host all for increased prodn. of polypeptides.
XX
XX Disclosure; Page 553; 34pp; Japanese.
XX
XX Sequence encodes beta-urogastrone under the control of a tac promoter.
XX The peptide may be expressed from plasmid pUGT 150s in a transformed
XX E.coli host. The plasmid may carry several separately expressing
XX sequences comprising a tac promoter, SD site, signal peptide, and coding
XX sequence, to produce beta-UG in high yield. (Updated on 10-MAR-2003 to
XX add missing OS field.) (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 1; Length 456;
Best Local Similarity 100.0%; Pred. No. 2.6e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTATTTAGAAAAATAAACAATAG 60
Db 173 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTATTTAGAAAAATAAACAATAG 114

Qy 61 GGGTTCGCGCACATTTCCCGGAAAGTGCACCTGACGTC 101
Db 113 GGGTTCGCGCACATTTCCCGGAAAGTGCACCTGACGTC 73

RESULT 8
AAN81765/c
ID AAN81765 standard; DNA; 456 BP.
XX
AC AAN81765;
XX
DT 25-MAR-2003 (revised)
DT 13-DEC-1990. (first entry)
```

```
AAN70833/c
ID AAN70833 standard; DNA; 456 BP.
XX
AC AAN70833;
XX
DT 25-MAR-2003 (revised)
DT 10-MAR-2003 (revised)
DT 18-JAN-1991 (first entry)
XX
DE Beta-urogastrone sequence.
XX
XX Tumour; inosine; DNA probe; ds.
XX
OS Unidentified.
XX
XX Key Location/Qualifiers
FT promoter 125..170
FT FT /*tag= b
FT RBS 200..204
FT FT /*tag= c
FT CDS 209..439
FT FT /*tag= a
FT FT sig_peptide 209..277
FT FT /*tag= d
XX
XX JP62244398-A.
XX
XX 24-OCT-1987.
XX
XX 16-APR-1986; 86JP-00087368.
XX
XX 16-APR-1986; 86JP-00087368.
XX
XX (SEKI ) SEKISUI CHEM IND CO LTD.
XX
XX WPI; 1987-339045/48.
XX
XX P-PSDB; AAP70505.
XX
XX Detection of DNA and/or RNA - by converting to single strand form and
XX using probe contg. labelled inosine deriv.
XX
XX Disclosure; Page 11; 11pp; Japanese.
XX
XX An example of a sequence detected by a probe consisting of polyinosine,
XX polydeoxyinosine, oligoinosine and/or oligodeoxyinosine is labeled. The
XX ssDNA and probe are hybridized and the existence of DNA in the product is
XX detected. It can be used to detect the presence of malignant tumour.
XX (Updated on 10-MAR-2003 to add missing OS field.) (Updated on 25-MAR-2003
XX to correct PA field.)
XX
SQ Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 1; Length 456;
Best Local Similarity 100.0%; Pred. No. 2.6e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTATTTAGAAAAATAAACAATAG 60
Db 173 AGGGTTATTGCTCATGCGGATACATATTTGAATGTTATTTAGAAAAATAAACAATAG 114

Qy 61 GGGTTCGCGCACATTTCCCGGAAAGTGCACCTGACGTC 101
Db 113 GGGTTCGCGCACATTTCCCGGAAAGTGCACCTGACGTC 73

RESULT 8
AAN81765/c
ID AAN81765 standard; DNA; 456 BP.
XX
AC AAN81765;
XX
DT 25-MAR-2003 (revised)
DT 13-DEC-1990. (first entry)
```

XX Sequence of HB101 (pGH201) encoding new beta-urogastrone deriv. Met (21),
DE Arg (53).
XX Gastric acid secretion; cell proliferation; hormone; ds.
XX Synthetic.
XX Key Location/Qualifiers
FH CDS 209..277
FT FT /*tag= a
FT CDS 278..439
FT FT /*tag= b
FT FT /product= "New beta-urogastrone deriv."
XX JP63012298-A.
XX 19-JAN-1988.
XX 30-JUN-1986; 86JP-00153783.
XX 30-JUN-1986; 86JP-00153783.
XX (EART) EARTH SEIYAKU KK.
XX WPI; 1988-054638/08.
XX P-PSDB; AAP81349.
XX New beta-urogastrone deriv. - has gastric acid secretion inhibition and
PT proliferation promotion activity.
XX Disclosure; Page 685; 76pp; Japanese.
XX The deriv. has various biological activities such as gastric acid
CC secretion inhibiting action, or cell proliferation promoting action. The
CC deriv. has the same biological or pharmacological activities as beta-
CC urogastrone. It is not susceptible to denaturation by oxidn. and is
CC chemically stable. Deriv. has resistance to proteolytic enzymes such as
CC pepsinase. (Updated on 25-MAR-2003 to correct PA field.)
XX Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 101; DB 1; Length 456;
Best Local Similarity 100.0%; Pred. No. 2.6e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGCTCATGCGGATACATATTGATGTTAGAAAAATAACAATAG 60
Db 173 AGGGTTATTGCTCATGCGGATACATATTGATGTTAGAAAAATAACAATAG 114
Qy 61 GGGTTCCGCGCACATTTCCCGAAAAAGTCCACCTGACGTC 101
Db 113 GGGTTCCGCGCACATTTCCCGAAAAAGTCCACCTGACGTC 73
RESULT 9
ABA90413/c
ID ABA90413 standard; DNA; 466 BP.
XX AC ABA90413;
XX 12-FEB-2002 (first entry)
XX Drosophila cell cycle progression protein coding sequence #48.
XX Antiproliferative; cytostatic; cardiant; immunosuppressive; meiosis;
KW antiinflammatory; antipsoriatic; dermatological; antifungal; mitosis;
KW antiparasitic; antimalarial; antirheumatic; antiarthritic; cell division;
KW cell cycle progression protein; tumour; proliferative disorder;
KW cardiovascular; autoimmune; dermatological disorder; ds.
XX Drosophila sp.
XX OS PI Fraser CW;
XX

PN WO200172774-A2.
XX 04-OCT-2001.
XX 23-MAR-2001; 2001WO-GB001297.
XX 24-MAR-2000; 2000GB-00007268.
XX (CYCL-) CYCLACEL LTD.
XX Deak P, Glover DM, Midgley C;
XX WPI; 2002-055132/07.
XX Polynucleotides encoding cell cycle progression proteins, useful for
PT treating a tumor or a proliferative disorder.
XX Claim 1; Page 99; 213pp; English.
XX The present invention relates to Drosophila cell cycle progression
CC proteins (AAM47572-AAM47608) and their coding sequences (ABA90366-
CC ABA90520). The coding sequences and proteins are useful for identifying a
CC substance capable of affecting the function of the corresponding gene, a
CC substance capable of inhibiting the cell division cycle, or capable of
CC inhibiting mitosis and/or meiosis. They can also be used in a method for
CC treating a tumour or proliferative disorder, cardiovascular disorders
CC (such as restenosis and cardiomyopathy), autoimmune disorders such as
CC (glomerulonephritis and rheumatoid arthritis), dermatological disorders
CC (such as psoriasis), antiinflammatory, antifungal and antiparasitic
CC disorders (such as malaria)
XX Sequence 466 BP; 135 A; 87 C; 93 G; 150 T; 0 U; 1 Other;
SQ
Query Match 100.0%; Score 101; DB 6; Length 466;
Best Local Similarity 100.0%; Pred. No. 2.6e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGCTCATGCGGATACATATTGATGTTAGAAAAATAACAATAG 60
Db 280 AGGGTTATTGCTCATGCGGATACATATTGATGTTAGAAAAATAACAATAG 221
Qy 61 GGGTTCCGCGCACATTTCCCGAAAAAGTCCACCTGACGTC 101
Db 220 GGGTTCCGCGCACATTTCCCGAAAAAGTCCACCTGACGTC 180
RESULT 10
AAX211173/c
ID AAX211173 standard; DNA; 487 BP.
XX AC AAX211173;
XX 05-MAY-1999 (first entry)
XX Polynucleotide sequence from the genome of Treponema pallidum.
XX Treponema pallidum infection; syphilis; Borrelia infection; animal;
KW enzyme production; ds.
XX Treponema pallidum.
XX WO9859034-A2.
XX 30-DEC-1998.
XX 23-JUN-1998; 98WO-US013041.
XX 24-JUN-1997; 97US-0050667P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX PI Fraser CW;
XX

PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241212P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.

PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-476223/51.
DR
XX
XX
PT
PT
PT
XX
PS
XX

Novel isolated prostate gland related polypeptide useful for diagnosis and treatment of disorders of prostate such as prostatodystonia, prostatosis, prostatitis, benign prostatic hypertrophy and malacoplakia.
Claim 1; SEQ ID NO 418; 512pp; English.

The invention relates to novel isolated prostate gland related nucleic acids (I) and polypeptides (II). (I) and (II) are useful for diagnosis, prognosis, prevention, and/or treatment of diseases and/or disorders of the prostate such as acute non-bacterial prostatitis, chronic non-bacterial prostatitis, acute bacterial prostatitis, prostatodystonia, prostatosis, granulomatous prostatitis, malacoplakia, benign prostatic hypertrophy or hyperplasia, and prostate neoplastic disorders, including adenocarcinomas, transitional cell carcinomas, ductal carcinomas, and squamous cell carcinomas. (I), (II) and antibody to (I) are useful for diagnosing and treating reproductive system disorders (Paget's disease), autoimmune disorders (systemic lupus erythematosus, rheumatoid arthritis), blood-related disorders (sickle cell anaemia), hyperproliferative disorders, urinary system disorders (glomerulonephritis), cardiovascular disorders (arrhythmias), respiratory disorders, musculoskeletal system disorders, neural activity and neurological disorders (Alzheimer's disease and Parkinson's disease), endocrine disorders (Addison's disease), gastrointestinal disorders (inflammatory disorders), liver disorders (biliary liver cirrhosis), pancreatic and gall bladder disorders, disorders of the large intestine, developmental and inherited disorders, diseases at the cellular level, and wound healing and epithelial cell proliferation. (I) or (II) is useful to prevent skin aging, for preventing hair loss, to maintain organs before transplantation, and as food additive or preservative.

Query Match 100.0%; Score 101; DB 4; Length 776;
Best Local Similarity 100.0%; Pred. No. 2.9e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGACGGGATACATATTGGAATGTTATTTAGAAAAATAACAATAG 60
|||
Db 546 AGGGTTATTGCTCATGACGGGATACATATTGGAATGTTATTTAGAAAAATAACAATAG 487
|||
Qy 61 GGTTTCGGCGCACATTTCCCGAAAGTGCACCTGACGTC 101
|||
Db 486 GGTTTCGGCGCACATTTCCCGAAAGTGCACCTGACGTC 446
|||

RESULT 15
AAS27819/c
ID AAS27819 standard; DNA; 776 BP.

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 05:15:57 ; Search time 961.667 Seconds
(without alignments)
3997.736 Million cell updates/sec

Title: US-09-482-682-65_COPY_8384_8484

Perfect score: 101
Sequence: 1 aggtttattgtctatgagc.....gaaagtgcacacgcagtc 101

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsl1:*
9: gb_gsl2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	101	100.0	300	4	BM078095
C 2	101	100.0	300	5	BU963956
C 3	101	100.0	300	5	BU964094
C 4	101	100.0	309	9	FR0009140
C 5	101	100.0	391	1	AL5971149
C 6	101	100.0	414	9	CC819240
C 7	101	100.0	417	4	BJ684174
C 8	101	100.0	491	9	CC819233
C 9	101	100.0	495	4	BI805285
C 10	101	100.0	495	9	CC818374
C 11	101	100.0	496	9	CC818523
C 12	101	100.0	503	9	CC819854
C 13	101	100.0	515	9	CC817752
C 14	101	100.0	518	9	CC817128
C 15	101	100.0	519	9	CC817162
C 16	101	100.0	519	9	CC817796
C 17	101	100.0	521	9	CC819067
C 18	101	100.0	533	9	CC819841
C 19	101	100.0	542	9	CC816892
C 20	101	100.0	550	7	CR766622
C 21	101	100.0	551	9	CC816905
C 22	101	100.0	554	9	CC819058
C 23	101	100.0	563	9	CC819270
C 24	101	100.0	566	9	CC816848

C 25	101	100.0	566	9	CC820024
C 26	101	100.0	567	9	CC817070
C 27	101	100.0	568	9	CC818640
C 28	101	100.0	571	9	CC816954
C 29	101	100.0	571	9	CC818423
C 30	101	100.0	580	9	CC819098
C 31	101	100.0	582	1	AL694813
C 32	101	100.0	583	9	CC817633
C 33	101	100.0	583	9	CC818436
C 34	101	100.0	586	9	CC816883
C 35	101	100.0	588	9	CC817788
C 36	101	100.0	588	9	CC818340
C 37	101	100.0	589	9	CC817595
C 38	101	100.0	590	9	CC819754
C 39	101	100.0	592	9	CC817679
C 40	101	100.0	592	9	CC818508
C 41	101	100.0	593	9	CC816942
C 42	101	100.0	593	9	CC817699
C 43	101	100.0	594	9	CC818287
C 44	101	100.0	594	9	CC818422
C 45	101	100.0	595	9	CC816929

ALIGNMENTS

RESULT 1
BM078095/c 300 bp mRNA linear EST 30-NOV-2001
LOCUS 83374 Hebeloma cylindrosporum functional cDNA library Hebeloma
DEFINITION 83374 Hebeloma cylindrosporum functional cDNA library Hebeloma
ACCESSION BM078095
VERSION BM078095.1 GI:17157967
KEYWORDS EST.
SOURCE Hebeloma cylindrosporum
ORGANISM Hebeloma cylindrosporum
REFERENCE 1 (bases 1 to 300)
AUTHORS Wipf D., Benjdia M., Tegeder M. and Frommer W.B.
TITLE Construction of a functional cDNA library from the ectomycorrhizal fungus Hebeloma cylindrosporum
JOURNAL Unpublished (2001)
COMMENT Contact: Wipf D
ZMBP - Center for Molecular Biology of Plants
University of Tuebingen
Auf der Morgenstelle 1, 72076 Tuebingen, Germany
Tel: 49 7071 2976160
Fax: 49 7071 293287
Email: daniel.wipf@zmbp.uni-tuebingen.de
PCR Primers
FORWARD: PDR196 5' primer (PMA 5')
HIGH quality sequence stop: 300
POLYA=NO.

FEATURES

Location/Qualifiers
1..300
/organism="Hebeloma cylindrosporum"
/mol_type="mRNA"
/strain="H1"
/db_xref="taxon:76867"
/tissue_type="Mycelia"
/lab_host="E. coli XL1-Blue"
/clone_lib="Hebeloma cylindrosporum functional cDNA library"
/note="vector: pDR 196 (unpublished); Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Query Match 100.0%; Score 101; DB 4; Length 300;
Best Local Similarity 100.0%; Pred. No. 8.1e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGTTATTGTCATGCGGATACATATTGTAATGTTAGAAAAATAACAAATAG 60

```

|||||
174 AGGTTATTGTCATGACGGATACATATTGATTTAGTATTAGAAAAATAACAATAG 115
|||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGACGTC 101
|||||
Db 114 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGACGTC 74
|||||

RESULT 2
BU963956/c
LOCUS
DEFINITION EST86 Hebeloma cylindrosporum functional cDNA library Hebeloma
cylindrosporum cDNA, mRNA sequence.
ACCESSION BU963956
VERSION BU963956.1 GI:24204753
KEYWORDS EST.
SOURCE Hebeloma cylindrosporum
ORGANISM Hebeloma cylindrosporum
Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
Agaricales; Cortinariaceae; Hebeloma.
REFERENCE 1 (bases 1 to 300)
AUTHORS Wipf,D., Benjdia,M., Rikirsch,E., Zimmermann,S., Tegeder,M. and
Frommer,W.B.
TITLE A Hebeloma cylindrosporum expression cDNA library for suppression
cloning in yeast mutants, complementation of a yeast his4 mutant
and EST analysis
JOURNAL Unpublished (2002)
COMMENT Contact: Wipf D
ZMBP - Center for Molecular Biology of Plants
University of Tuebingen
Auf der Morgenstelle 1, 72076 Tuebingen, Germany
Tel: 49 7071 2976160
Fax: 49 7071 293287
Email: daniel.wipf@zmbp.uni-tuebingen.de
homolog to Enterobacteria phage f1 ; ampicillinase (1e-10) .
FEATURES
source
1..300
/organism="Hebeloma cylindrosporum"
/mol_type="mRNA"
/strain="H1"
/db_xref="taxon:76867"
/tissue_type="Mycelia"
/lab_host="E. coli XL1-Blue"
/clone_lib="Hebeloma cylindrosporum functional cDNA
library"
/notes="Vector: pDR 196 (unpublished); Site_1: EcoRI;
Site_2: XhoI"

Query Match 100.0%; Score 101; DB 5; Length 300;
Best Local Similarity 100.0%; Pred. No. 8.1e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGTTATTGTCATGACGGATACATATTGATTTAGTATTAGAAAAATAACAATAG 60
|||||
Db 170 AGGTTATTGTCATGACGGATACATATTGATTTAGTATTAGAAAAATAACAATAG 111
|||||

RESULT 3
BU964094/c
LOCUS
DEFINITION EST226 Hebeloma cylindrosporum functional cDNA library Hebeloma
cylindrosporum cDNA, mRNA sequence.
ACCESSION BU964094
VERSION BU964094.1 GI:24204891
KEYWORDS EST.
SOURCE Hebeloma cylindrosporum
ORGANISM Hebeloma cylindrosporum
Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;

```

```

Agaricales; Cortinariaceae; Hebeloma.
REFERENCE 1 (bases 1 to 300)
AUTHORS Wipf,D., Benjdia,M., Rikirsch,E., Zimmermann,S., Tegeder,M. and
Frommer,W.B.
TITLE A Hebeloma cylindrosporum expression cDNA library for suppression
cloning in yeast mutants, complementation of a yeast his4 mutant
and EST analysis
JOURNAL Unpublished (2002)
COMMENT Contact: Wipf D
ZMBP - Center for Molecular Biology of Plants
University of Tuebingen
Auf der Morgenstelle 1, 72076 Tuebingen, Germany
Tel: 49 7071 2976160
Fax: 49 7071 293287
Email: daniel.wipf@zmbp.uni-tuebingen.de
no homolog below 1e-10.
FEATURES
Location/Qualifiers
source
1..300
/organism="Hebeloma cylindrosporum"
/mol_type="mRNA"
/strain="H1"
/db_xref="taxon:76867"
/tissue_type="Mycelia"
/lab_host="E. coli XL1-Blue"
/clone_lib="Hebeloma cylindrosporum functional cDNA
library"
/notes="Vector: pDR 196 (unpublished); Site_1: EcoRI;
Site_2: XhoI"

Query Match 100.0%; Score 101; DB 5; Length 300;
Best Local Similarity 100.0%; Pred. No. 8.1e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGTTATTGTCATGACGGATACATATTGATTTAGTATTAGAAAAATAACAATAG 60
|||||
Db 170 AGGTTATTGTCATGACGGATACATATTGATTTAGTATTAGAAAAATAACAATAG 111
|||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGACGTC 101
|||||
Db 110 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGACGTC 70
|||||

RESULT 4
FR0009140
LOCUS
DEFINITION F.rubripes GSS sequence, clone 010H20aC4, genomic survey sequence.
ACCESSION AL000426
VERSION AL000426.1 GI:2438278
KEYWORDS GSS; genome survey sequence.
SOURCE Takifugu rubripes (Fugu rubripes)
ORGANISM Takifugu rubripes
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Takifugu.
REFERENCE 1
AUTHORS Elgar,G., Clark,M.S., Meek,S., Smith,S., Warner,S., Edwards,Y.J.,
Bouchireh,N., Cottage,A., Yeo,G.S., Umrانيا,Y., Williams,G. and
Brenner,S.
TITLE Generation and analysis of 25 Mb of genomic DNA from the pufferfish
Fugu rubripes by sequence scanning
JOURNAL Genome Res. 9 (10), 960-971 (1999)
MEDLINE 99455097
PUBMED 10523524
REFERENCE 2 (bases 1 to 309)
AUTHORS Elgar,G., Clark,M., Smith,S., Meek,S., Warner,S., Umrانيا,Y.,
Williams,G. and Brenner,S.
TITLE Direct Submission
JOURNAL Submitted (09-SEP-1997) MRC Human Genome Mapping Project Resource
Centre Hinxton, Cambridge, CB10 1SB. Email: biohelp@hgm.mrc.ac.uk
COMMENT Vector: pBluescript II KS
V_type: phagemid

```

```

PRIMER: KS
DESCR:
One pass dye-terminator sequencing of cosmid cloned genomic
sequence.
FEATURES
    source
        Location/Qualifiers
            1..309
                /organism="Takifugu rubripes"
                /mol_type="genomic DNA"
                /db_xref="taxon:31033"
                /clone="010H20aC4"
                /clone_lib="cosmid 010H20"
ORIGIN
Query Match      100.0%; Score 101; DB 9; Length 309;
Best Local Similarity 100.0%; Pred. No. 8.2e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGTTATTGCTCATGCGGATACATATTTGAATGTTATTTAGAAAAATAAACAATAG 60
Db 39 AGGTTATTGCTCATGCGGATACATATTTGAATGTTATTTAGAAAAATAAACAATAG 98
QY 61 GGGTTCGGCGACATTTCCCGAAAAGTGCCACCTGACGTC 101
Db 99 GGGTTCGGCGACATTTCCCGAAAAGTGCCACCTGACGTC 139
RESULT 5
LOCUS
DEFINITION
AL597149 391 bp mRNA linear EST 04-SEP-2003
DFK2p313J1611_r1 313 (synonym: hlcc2) Homo sapiens cDNA clone
DFK2p313J1611_5', mRNA sequence.
ACCESSION
AL597149
VERSION
AL597149.1 GI:15154845
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 391)
Koehler, K., Beyer, A., Mewes, H.W., Weil, B. and Wiemann, S.)
EST (Koehler, K., Beyer, A., Mewes, H.W., Weil, B. and Wiemann, S.)
Unpublished (1999)
Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by BMFZ (Biomedical Research Center at the Charite,
Berlin/Germany) within the cDNA sequencing consortium of the German
Genome Project.
No sl sequence available.
This clone (DKF2p313J1611) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
    source
        Location/Qualifiers
            1..391
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="DKF2p313J1611"
                /dev_stage="adult"
                /lab_host="DH10B"
                /clone_lib="313 (synonym: hlcc2)"
                /note="vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB;
                cDNA-collection"
ORIGIN
Query Match      100.0%; Score 101; DB 1; Length 391;
Best Local Similarity 100.0%; Pred. No. 8.3e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGTTATTGCTCATGCGGATACATATTTGAATGTTATTTAGAAAAATAAACAATAG 60

```

```

Db 228 AGGTTATTGCTCATGCGGATACATATTTGAATGTTATTTAGAAAAATAAACAATAG 287
QY 61 GGGTTCGGCGACATTTCCCGAAAAGTGCCACCTGACGTC 101
Db 288 GGGTTCGGCGACATTTCCCGAAAAGTGCCACCTGACGTC 328
RESULT 6
LOCUS
DEFINITION
CC819240 414 bp DNA linear GSS 17-JUL-2003
100005D19R Oxytricha plasmid UUGC10 library Sterkiella
histriomuscorum genomic clone UUGC100005D19 R, genomic survey
sequence.
ACCESSION
CC819240
VERSION
CC819240.1 GI:32899308
KEYWORDS
GSS.
SOURCE
Sterkiella histriomuscorum (Oxytricha trifallax)
ORGANISM
Sterkiella histriomuscorum
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
Stichotrichida; Oxytrichidae; Sterkiella.
1 (bases 1 to 414)
Dunn, D., Doak, T., Herrick, G. and Weiss, R.
Paired end reads from plasmid inserts of Oxytricha trifallax
macronuclear chromosomes
Unpublished (2003)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Plate: 0005 row: D column: 19
Seq primer: CACACAGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 414.
FEATURES
    source
        Location/Qualifiers
            1..414
                /organism="Sterkiella histriomuscorum"
                /mol_type="genomic DNA"
                /db_xref="taxon:94289"
                /clone="UUGC100005D19"
                /lab_host="E. Coli strain XL10-Gold, T4-resistant, F-"
                /clone_lib="Oxytricha plasmid UUGC10 library"
                /note="Vector: FWD42nv; Purified macronuclear chromosomal
                DNA from Oxytricha trifallax was blunt end-repaired with
                T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
                oligonucleotides were ligated to the blunt ends in high
                molar excess. Vector DNA was prepared from a derivative of
                pWD42 (GI|4732114|gb|AF129072.1), a copy-number inducible
                derivative of plasmid Ri. The vector was ligated with
                adaptors complementary to the insert adaptors and
                purified. The sheared, adapted mouse DNA was annealed to
                adapted vector DNA, and transformed into
                chemically-competent E. Coli XL10-Gold (Stratagene) cells
                and selected for ampicillin resistance."
ORIGIN
Query Match      100.0%; Score 101; DB 9; Length 414;
Best Local Similarity 100.0%; Pred. No. 8.3e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGTTATTGCTCATGCGGATACATATTTGAATGTTATTTAGAAAAATAAACAATAG 60
Db 414 AGGTTATTGCTCATGCGGATACATATTTGAATGTTATTTAGAAAAATAAACAATAG 355
QY 61 GGGTTCGGCGACATTTCCCGAAAAGTGCCACCTGACGTC 101
Db 354 GGGTTCGGCGACATTTCCCGAAAAGTGCCACCTGACGTC 314

```

```

RESULT 7
BJ684174/c
LOCUS BJ684174 HBST library Haplochromis chilotes cDNA clone no90c12, 417 bp mRNA linear EST 23-APR-2004
DEFINITION BJ684174 HBST library Haplochromis chilotes cDNA clone no90c12, mRNA sequence.
ACCESSION BJ684174
VERSION BJ684174.1 GI:46527295
KEYWORDS EST.
SOURCE Haplochromis chilotes
ORGANISM Haplochromis chilotes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes;
Labroides; Cichlidae; Haplochromis.
1 (bases 1 to 417)
Watanabe,M., Kobayashi,N., Shin-i,T., Kohara,Y. and Okada,N.
Orig sequences of cichlid in Lake Victoria are essentially same
Unpublished (2004)
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1. .417
/organism="Haplochromis chilotes"
/mol_type="mRNA"
/db_xref="taxon:257977"
/clone="no90c12"
/tissue_type="jaw"
/dev_stage="varied"
/clone_lib="HCEST library"

ORIGIN
Query Match 100.0%; Score 101; DB 4; Length 417;
Best Local Similarity 100.0%; Pred. No. 8.3e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGTTATTGTCATGAGCGGATACATATTTGAATGTTATTTAGAAAAATAACAATAG 60
|||||
Db 129 AGGTTATTGTCATGAGCGGATACATATTTGAATGTTATTTAGAAAAATAACAATAG 70
|||||

Qy 61 GGGTTCGGCGACATTTCCCGAAAAGTGCACCTGACGTC 101
|||||
Db 69 GGGTTCGGCGACATTTCCCGAAAAGTGCACCTGACGTC 29
|||||

RESULT 8
CC819923/c
LOCUS CC819923 491 bp DNA linear GSS 17-JUL-2003
DEFINITION 10006J13R Oxytricha plasmid UUGC10 library Sterkiella histriomuscorum genomic clone UUGC10006J13 R, genomic survey sequence.
ACCESSION CC819923
VERSION CC819923.1 GI:32900671
KEYWORDS GSS.
SOURCE Sterkiella histriomuscorum (Oxytricha trifallax)
ORGANISM Sterkiella histriomuscorum
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
Stichotrichida; Oxytrichidae; Sterkiella.
1 (bases 1 to 491)
Dunn,D., Doak,T., Herrick,G. and Weiss,R.
Paired end reads from plasmid inserts of Oxytricha trifallax
macronuclear chromosomes
Unpublished (2003)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606

```

```

Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Plate: 0006 row: J column: 13
Seq primer: CACACAGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 491.
Location/Qualifiers
1. .491
/organism="Sterkiella histriomuscorum"
/mol_type="genomic DNA"
/db_xref="taxon:94289"
/clone="UUGC10006J13"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Oxytricha plasmid UUGC10 library"
/notes="Vector: PWD42nv; Purified macronuclear chromosomal DNA from Oxytricha trifallax was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. Coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN
Query Match 100.0%; Score 101; DB 9; Length 491;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGTTATTGTCATGAGCGGATACATATTTGAATGTTATTTAGAAAAATAACAATAG 60
|||||
Db 412 AGGTTATTGTCATGAGCGGATACATATTTGAATGTTATTTAGAAAAATAACAATAG 353
|||||

Qy 61 GGGTTCGGCGACATTTCCCGAAAAGTGCACCTGACGTC 101
|||||
Db 352 GGGTTCGGCGACATTTCCCGAAAAGTGCACCTGACGTC 312
|||||

RESULT 9
BI805285
LOCUS BI805285 495 bp mRNA linear EST 02-OCT-2001
DEFINITION S035A01 Stem library from Oryza sativa (3-5 leaf stage) Oryza sativa cDNA clone S035A01, mRNA sequence.
ACCESSION BI805285
VERSION BI805285.1 GI:15852489
KEYWORDS EST.
SOURCE Oryza sativa
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 495)
Dong,H.T., Li,D.B., Zhuang,X.F., Dai,C.G., Sun,L.X., Pei,Y.X.,
Wu,H.F., Jiang,Y.X., Yu,F.C., Gao,Q.K. and Lou,Y.C.
A Gene Expression Screen in Oryza sativa
Unpublished (2001)
Contact: Haitao Dong, Debao Li
Bioinformatics and Gene Network Research Group
Zhejiang University
Kaixuan Road 268#, Hangzhou, Zhejiang, P.R.China
Tel: 0086-571-86892051
Fax: 0086-571-86961525
Email: webmaster@estarray.org, URL: http://www.estarray.org
Seq primer: M13 forward primer.
Location/Qualifiers
1. .495
/organism="Oryza sativa"
/mol_type="mRNA"
/db_xref="taxon:4530"
/clone="S035A01"

FEATURES
source

```

```

/tissue_type="Stem"
/dev stage="3-5 leaf stage"
/clone_lib="Stem library from Oryza sativa (3-5 leaf
stage)"
/note="vector: pSport2"

ORIGIN
Query Match      100.0%; Score 101; DB 4; Length 495;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAATAG 60
    |||||
Db 62 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAATAG 121

Qy 61 GGGTTCGCGCACATTTCCTCCGAAAAGTGCACCTGACGTC 101
    |||||
Db 122 GGGTTCGCGCACATTTCCTCCGAAAAGTGCACCTGACGTC 162

RESULT 10
CC818374/c
LOCUS      CC818374      495 bp      DNA      linear      GSS 17-JUL-2003
DEFINITION 100004807R Oxytricha plasmid UUGC10 library Sterkiella
histriomuscorum genomic clone UUGC100004B07 R, genomic survey
sequence.
ACCESSION  CC818374
VERSION     CC818374.1 GI:32897661
KEYWORDS   GSS.
SOURCE     Sterkiella histriomuscorum (Oxytricha trifallax)
ORGANISM   Sterkiella histriomuscorum
            Sterkiella histriomuscorum
            Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
            Stichotrichida; Oxytrichidae; Sterkiella.
REFERENCE  1 (bases 1 to 495)
AUTHORS   Dunn,D., Doak,T., Herrick,G. and Weiss,R.
TITLE     Paired end reads from plasmid inserts of Oxytricha trifallax
          macronuclear chromosomes
JOURNAL   Unpublished (2003)
COMMENT   Contact: Robert B. Weiss
          University of Utah Genome Center
          University of Utah
          Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
          84112, USA
          Tel: 801 585 5606
          Fax: 801 585 7177
          Email: ddunn@genetics.utah.edu
          Plate: 0004 row: B column: 07
          Seq primer: CACACAGGAAACAGCTATGACC
          Class: plasmid ends
          High quality sequence stop: 495.
FEATURES   source
            1..495
             /organism="Sterkiella histriomuscorum"
             /mol_type="genomic DNA"
             /db_xref="taxon:94289"
             /clone="UUGC100004B07"
             /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
             /note="vector: PWD42nv; Purified macronuclear chromosomal
             DNA from Oxytricha trifallax was blunt end-repaired with
             T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
             oligonucleotides were ligated to the blunt ends in high
             molar excess. Vector DNA was prepared from a derivative of
             PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
             derivative of plasmid R1. The vector was ligated with
             adaptors complementary to the insert adaptors and
             adapted mouse DNA, and transformed into
             chemically-competent E. Coli XL10-Gold (Stratagene) cells
             and selected for ampicillin resistance."

ORIGIN
Query Match      100.0%; Score 101; DB 9; Length 495;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAATAG 60
    |||||
Db 391 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAATAG 332

Qy 61 GGGTTCGCGCACATTTCCTCCGAAAAGTGCACCTGACGTC 101
    |||||

```

```

Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAATAG 60
    |||||
Db 392 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAATAG 333

Qy 61 GGGTTCGCGCACATTTCCTCCGAAAAGTGCACCTGACGTC 101
    |||||
Db 332 GGGTTCGCGCACATTTCCTCCGAAAAGTGCACCTGACGTC 292

RESULT 11
CC818523/c
LOCUS      CC818523      496 bp      DNA      linear      GSS 17-JUL-2003
DEFINITION 100004L13R Oxytricha plasmid UUGC10 library Sterkiella
histriomuscorum genomic clone UUGC100004L13 R, genomic survey
sequence.
ACCESSION  CC818523
VERSION     CC818523.1 GI:32897943
KEYWORDS   GSS.
SOURCE     Sterkiella histriomuscorum (Oxytricha trifallax)
ORGANISM   Sterkiella histriomuscorum
            Sterkiella histriomuscorum
            Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
            Stichotrichida; Oxytrichidae; Sterkiella.
REFERENCE  1 (bases 1 to 496)
AUTHORS   Dunn,D., Doak,T., Herrick,G. and Weiss,R.
TITLE     Paired end reads from plasmid inserts of Oxytricha trifallax
          macronuclear chromosomes
JOURNAL   Unpublished (2003)
COMMENT   Contact: Robert B. Weiss
          University of Utah Genome Center
          University of Utah
          Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
          84112, USA
          Tel: 801 585 5606
          Fax: 801 585 7177
          Email: ddunn@genetics.utah.edu
          Plate: 0004 row: L column: 13
          Seq primer: CACACAGGAAACAGCTATGACC
          Class: plasmid ends
          High quality sequence stop: 496.
FEATURES   source
            1..496
             /organism="Sterkiella histriomuscorum"
             /mol_type="genomic DNA"
             /db_xref="taxon:94289"
             /clone="UUGC100004L13"
             /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
             /note="vector: PWD42nv; Purified macronuclear chromosomal
             DNA from Oxytricha trifallax was blunt end-repaired with
             T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
             oligonucleotides were ligated to the blunt ends in high
             molar excess. Vector DNA was prepared from a derivative of
             PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
             derivative of plasmid R1. The vector was ligated with
             adaptors complementary to the insert adaptors and
             purified. The sheared, adapted mouse DNA was annealed to
             adapted vector DNA, and transformed into
             chemically-competent E. Coli XL10-Gold (Stratagene) cells
             and selected for ampicillin resistance."

ORIGIN
Query Match      100.0%; Score 101; DB 9; Length 496;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAATAG 60
    |||||
Db 391 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAATAG 332

Qy 61 GGGTTCGCGCACATTTCCTCCGAAAAGTGCACCTGACGTC 101
    |||||

```

```

Db      331 GGGTTCGGCGACATTTCCCGAAAGTGCCACCTGACGTC 291
|||||
CC8119854      503 bp      DNA      linear      GSS 17-JUL-2003
100006N08R Oxytricha plasmid UUGC10 library Sterkiella
histriomuscorum genomic clone UUGC100006N08 R, genomic survey
sequence.
ACCESSION      CC8119854
VERSION        CC8119854.1 GI:32900533
KEYWORDS
SOURCE
ORGANISM      Sterkiella histriomuscorum (Oxytricha trifallax)
Sterkiella histriomuscorum
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
Stichotrichida; Oxytrichidae; Sterkiella.
REFERENCE      1 (bases 1 to 503)
AUTHORS      Dunn,D., Doak,T., Herrick,G. and Weiss,R.
TITLE      Paired end reads from plasmid inserts of Oxytricha trifallax
macronuclear chromosomes
JOURNAL
COMMENT      Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Plate: 0006 row: N column: 08
Seq primer: CACACAGGAACACGCTATGACC
Class: plasmid ends
High quality sequence stop: 503.
FEATURES
source
1..503
/organism="Sterkiella histriomuscorum"
/mol_type="genomic DNA"
/db_xref="taxon:94289"
/clone="UUGC100006N08"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Oxytricha plasmid UUGC10 library"
/note="Vector: PWD42nv; Purified macronuclear chromosomal
DNA from Oxytricha trifallax was blunt end-repaired with
T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
oligonucleotides were ligated to the blunt ends in high
molar excess. Vector DNA was prepared from a derivative of
PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
derivative of plasmid R1. The vector was ligated with
adaptors complementary to the insert adaptors and
purified. The sheared, adaptor mouse DNA was annealed to
adaptor vector DNA, and transformed into
chemically-competent E. Coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
ORIGIN
Query Match      100.0%; Score 101; DB 9; Length 503;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1 AGGGTTATTGTCATGCGGATACATATTGTAATGTTATTAGAAAAATAACAATAG 60
|||||
Db      410 AGGGTTATTGTCATGCGGATACATATTGTAATGTTATTAGAAAAATAACAATAG 351
|||||
Qy      61 GGGTTCGGCGACATTTCCCGAAAGTGCCACCTGACGTC 101
|||||
Db      350 GGGTTCGGCGACATTTCCCGAAAGTGCCACCTGACGTC 310
|||||
RESULT 12
CC8119854/c
LOCUS
DEFINITION      100006N08R Oxytricha plasmid UUGC10 library Sterkiella
histriomuscorum genomic clone UUGC100006N08 R, genomic survey
sequence.
ACCESSION      CC8119854
VERSION        CC8119854.1 GI:32900533
KEYWORDS
SOURCE
ORGANISM      Sterkiella histriomuscorum (Oxytricha trifallax)
Sterkiella histriomuscorum
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
Stichotrichida; Oxytrichidae; Sterkiella.
REFERENCE      1 (bases 1 to 503)
AUTHORS      Dunn,D., Doak,T., Herrick,G. and Weiss,R.
TITLE      Paired end reads from plasmid inserts of Oxytricha trifallax
macronuclear chromosomes
JOURNAL
COMMENT      Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Plate: 0006 row: N column: 08
Seq primer: CACACAGGAACACGCTATGACC
Class: plasmid ends
High quality sequence stop: 503.
FEATURES
source
1..503
/organism="Sterkiella histriomuscorum"
/mol_type="genomic DNA"
/db_xref="taxon:94289"
/clone="UUGC100006N08"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Oxytricha plasmid UUGC10 library"
/note="Vector: PWD42nv; Purified macronuclear chromosomal
DNA from Oxytricha trifallax was blunt end-repaired with
T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
oligonucleotides were ligated to the blunt ends in high
molar excess. Vector DNA was prepared from a derivative of
PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
derivative of plasmid R1. The vector was ligated with
adaptors complementary to the insert adaptors and
purified. The sheared, adaptor mouse DNA was annealed to
adaptor vector DNA, and transformed into
chemically-competent E. Coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
ORIGIN
Query Match      100.0%; Score 101; DB 9; Length 503;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1 AGGGTTATTGTCATGCGGATACATATTGTAATGTTATTAGAAAAATAACAATAG 60
|||||
Db      410 AGGGTTATTGTCATGCGGATACATATTGTAATGTTATTAGAAAAATAACAATAG 351
|||||
Qy      61 GGGTTCGGCGACATTTCCCGAAAGTGCCACCTGACGTC 101
|||||
Db      350 GGGTTCGGCGACATTTCCCGAAAGTGCCACCTGACGTC 310
|||||
RESULT 13
CC8117752/c
LOCUS
DEFINITION      100003C16R Oxytricha plasmid UUGC10 library Sterkiella
histriomuscorum genomic clone UUGC100003C16 R, genomic survey
sequence.
ACCESSION      CC8117752
VERSION        CC8117752.1 GI:328997039
KEYWORDS
SOURCE
ORGANISM      Sterkiella histriomuscorum (Oxytricha trifallax)
Sterkiella histriomuscorum
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
Stichotrichida; Oxytrichidae; Sterkiella.
REFERENCE      1 (bases 1 to 515)
AUTHORS      Dunn,D., Doak,T., Herrick,G. and Weiss,R.
TITLE      Paired end reads from plasmid inserts of Oxytricha trifallax
macronuclear chromosomes
JOURNAL
COMMENT      Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Plate: 0003 row: C column: 16
Seq primer: CACACAGGAACACGCTATGACC
Class: plasmid ends
High quality sequence stop: 515.
FEATURES
source
1..515
/organism="Sterkiella histriomuscorum"
/mol_type="genomic DNA"
/db_xref="taxon:94289"
/clone="UUGC100003C16"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Oxytricha plasmid UUGC10 library"
/note="Vector: PWD42nv; Purified macronuclear chromosomal
DNA from Oxytricha trifallax was blunt end-repaired with
T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
oligonucleotides were ligated to the blunt ends in high
molar excess. Vector DNA was prepared from a derivative of
PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
derivative of plasmid R1. The vector was ligated with
adaptors complementary to the insert adaptors and
purified. The sheared, adaptor mouse DNA was annealed to
adaptor vector DNA, and transformed into
chemically-competent E. Coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
ORIGIN
Query Match      100.0%; Score 101; DB 9; Length 515;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1 AGGGTTATTGTCATGCGGATACATATTGTAATGTTATTAGAAAAATAACAATAG 60
|||||
Db      412 AGGGTTATTGTCATGCGGATACATATTGTAATGTTATTAGAAAAATAACAATAG 353
|||||
Qy      61 GGGTTCGGCGACATTTCCCGAAAGTGCCACCTGACGTC 101
|||||
Db      352 GGGTTCGGCGACATTTCCCGAAAGTGCCACCTGACGTC 312
|||||
RESULT 14
CC8117128/c
LOCUS
DEFINITION      100002D21R Oxytricha plasmid UUGC10 library Sterkiella
histriomuscorum genomic clone UUGC100002D21 R, genomic survey
sequence.
ACCESSION      CC8117128
VERSION        CC8117128.1 GI:328996415
KEYWORDS
SOURCE
ORGANISM      Sterkiella histriomuscorum (Oxytricha trifallax)
Sterkiella histriomuscorum
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;

```

University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunne@genetics.utah.edu
Place: 0002 row: J column: 19
Seq primer: CACACAGGAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 519.

FEATURES

```

585 7177
hum@genetics.utah.edu
302 row: D column: 21
accr: CACACGGAACAGCTATGACC
asmid ends
ity sequence stop: 518.
Location/Qualifiers
1. 518
/organism="Sterkiella histriomuscorum"
/mol_type="genomic DNA"
/db_xref="taxon:94289"
/clone="UUGC10002021"
/lab_host="E. Coli strain X110-Gold, Tl-resistant, F."
/clone_lib="Oxyricha plasmid UUGC10 library"
/notes="vector: PWD42nv; Purified macronuclear
DNA from Oxyricha trifallax was blunt end-repaired with
T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
oligonucleotides were ligated to the blunt ends in high

```

ORIGIN

derivative of plasmid RI. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. Coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

RESULT 15
CC817162/c
LOCUS
CC817162 519 bp DNA linear GSS 17-JUL-2003
DEFINITION
100002J19R Oxytricha plasmid UUC10 library Sterkiella
histrionuscorum genomic clone UUGC100002J19 R, genomic survey
sequence.

sequence.
CC817162
ACCESSION
VERSION
CC817162.1
Gr:32896449
KEYWORDS
GSS.

REINFORCED SOURCE	ORGANISM
USSR	<i>Sterkiella histriomuscorum</i> (<i>Oxytricha trifallax</i>)
	<i>Sterkiella histriomuscorum</i>
	Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia; Stichotrichida; Oxytrichidae; <i>Sterkiella</i> .

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1. (bases 1 to 519)
Oxytrichinae; Oxytrichidae; Sterkiella.
Dunn, D., Doak, T., Herrick, G. and Weiss, R.
Paired end reads from plasmid inserts of *Oxytricha trifallax*
macronuclear chromosomes
Unpublished (2003)
Contact: Robert B. Weiss
University of Utah Genome Center

UNPUBLISHED (2003)
Contact: Robert B. Weiss
University of Utah Genome Center

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:39:07 ; Search time 749.127 Seconds
(without alignments)
6468.225 Million cell updates/sec

Title: US-09-482-682-8_COPY_1_100
Perfect score: 100
Sequence: 1 gacggatcggagatctccc.....ctgtccctgtgtgtgtt 100

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_on.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100	100.0	3853	6 AR098190	AR098190 Sequence
2	100	100.0	3853	6 AR207832	AR207832 Sequence
3	100	100.0	3853	6 BD009729	BD009729 Tissue sp
4	100	100.0	3986	12 PCDNA32EO	X90639 Cloning vec
5	100	100.0	4026	6 AR098191	AR098191 Sequence
6	100	100.0	4026	6 AR207833	AR207833 Sequence
7	100	100.0	4026	6 BD009730	BD009730 Tissue sp
8	100	100.0	4249	6 AR098192	AR098192 Sequence
9	100	100.0	4249	6 AR207834	AR207834 Sequence
10	100	100.0	4249	6 BD009731	BD009731 Tissue sp
11	100	100.0	4341	6 A38214	A38214 Sequence 58
12	100	100.0	4341	6 AX286570	AX286570 Sequence
13	100	100.0	4597	6 AX060344	AX060344 Sequence
14	100	100.0	4840	6 AX133940	AX133940 Sequence
15	100	100.0	5053	6 BD238492	BD238492 Expressio
16	100	100.0	5070	6 AX234391	AX234391 Sequence
17	100	100.0	5082	6 A91754	A91754 Sequence 10
18	100	100.0	5082	6 BD085110	BD085110 Vertebrat
19	100	100.0	5162	6 AX951626	AX951626 Sequence

20	100	100.0	5257	12 CV089673	U89673 Cloning vec
21	100	100.0	5432	6 BD234590	BD234590 Screening
22	100	100.0	5432	6 AX026821	AX026821 Sequence
23	100	100.0	5446	6 AX319694	AX319694 Sequence
24	100	100.0	5618	6 A44171	A44171 Sequence 1
25	100	100.0	5618	6 AR116416	AR116416 Sequence
26	100	100.0	5618	6 AR222266	AR222266 Sequence
27	100	100.0	5618	6 AR411127	AR411127 Sequence
28	100	100.0	5639	12 AY437643	AY437643 Expressio
29	100	100.0	5651	6 AX211282	AX211282 Sequence
30	100	100.0	5651	6 AX349366	AX349366 Sequence
31	100	100.0	5653	6 I56772	I56772 Sequence 3
32	100	100.0	5653	6 I95540	I95540 Sequence 1
33	100	100.0	5726	12 CV089672	U89672 Cloning vec
34	100	100.0	5731	6 AX202478	AX202478 Sequence
35	100	100.0	5900	6 AX573107	AX573107 Sequence
36	100	100.0	5995	6 AX685746	AX685746 Sequence
37	100	100.0	6090	6 A63067	A63067 Sequence 11
38	100	100.0	6148	6 BD181637	BD181637 Novel mel
39	100	100.0	6148	6 AX342685	AX342685 Sequence
40	100	100.0	6149	6 BD181638	BD181638 Novel mel
41	100	100.0	6149	6 AX342686	AX342686 Sequence
42	100	100.0	6180	6 AX207724	AX207724 Sequence
43	100	100.0	6186	6 AX211281	AX211281 Sequence
44	100	100.0	6186	6 AX349365	AX349365 Sequence
45	100	100.0	6200	6 BD232461	BD232461 Recombina

ALIGNMENTS

RESULT 1	AR098190	Sequence 5 from patent US 6074850.	3853 bp	DNA	linear	PAT 14-FEB-2001
LOCUS	AR098190					
DEFINITION	AR098190					
ACCESSION	AR098190.1	GI:12807447				
VERSION	AR098190.1					
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unclassified.					
REFERENCE	1 (bases 1 to 3853)					
AUTHORS	Antelman,D., Gregory,R.J. and Wills,K.N.					
TITLE	Retinoblastoma fusion polypeptides					
JOURNAL	Patent: US 6074850-A 5 13-JUN-2000;					
FEATURES	Location/Qualifiers					
source	1..3853					
	/organism="unknown"					
	/mol_type="unassigned DNA"					
ORIGIN						

Query Match	100.0%;	Score 100;	DB 6;	Length 3853;		
Best Local Similarity	100.0%;	Pred. No. 9.4e-24;				
Matches 100;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;		
Qy	1	GACGGATCGGAGATCTCCCGATCCCGATCGCTCTCAGTACAACTGCTCTGATG	60			
Db	1	GACGGATCGGAGATCTCCCGATCCCGATCGCTCTCAGTACAACTGCTCTGATG	60			
Qy	61	CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT	100			
Db	61	CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT	100			
RESULT 2	AR207832	Sequence 5 from patent US 6379927.	3853 bp	DNA	linear	PAT 20-JUN-2002
LOCUS	AR207832					
DEFINITION	AR207832					
ACCESSION	AR207832.1	GI:21507688				
VERSION	AR207832.1					
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unknown.					

```
Unclassified.
1 (bases 1 to 3853)
Antelman,D., Gregory,R.J. and Wills,K.N.
Retinoblastoma fusion proteins
Patent: US 6379927-A 5 30-APR-2002;
Location/Qualifiers
1. 3853
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 3853;
Best Local Similarity 100.0%; Pred. No. 9.4e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
|
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
|
Qy 61 CCGCATAGTTAAGCAGTATCTGCTCCCTGCTTGTGTGTT 100
|
Db 61 CCGCATAGTTAAGCAGTATCTGCTCCCTGCTTGTGTGTT 100
|

RESULT 3
BD009729
LOCUS 3853 bp DNA linear PAT 31-JAN-2002
DEFINITION Tissue specific expression of retinoblastoma protein.
ACCESSION BD009729
VERSION BD009729.1 GI:18638102
KEYWORDS JP 2001503638-A/3.
SOURCE unidentified
ORGANISM unidentified
unclassified.

REFERENCE 1 (bases 1 to 3853)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Tissue specific expression of retinoblastoma protein
JOURNAL Patent: JP 2001503638-A 3 21-MAR-2001;
COMMENT CANJI INC
OS Unidentified
PN JP 2001503638-A/3
PD 21-MAR-2001
PF 13-NOV-1997 JP 1998522958
PR 15-NOV-1996 US 08/751517,14-FEB-1997 US 08/801092 PT
DOUGLAS ANTELMAN,RICHARD J GREGORY,KENNETH N WILLS PC
C07H21/04,C07K5/00,A61K38/00,A61K35/12
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT CDS 209..862.
Location/Qualifiers
1. 3853
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

FEATURES
source
ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 3853;
Best Local Similarity 100.0%; Pred. No. 9.4e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
|
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
|
Qy 61 CCGCATAGTTAAGCAGTATCTGCTCCCTGCTTGTGTGTT 100
|
Db 61 CCGCATAGTTAAGCAGTATCTGCTCCCTGCTTGTGTGTT 100
|

RESULT 4
PCDNA3ZEO
LOCUS 3986 bp DNA linear SYN 16-AUG-1995
DEFINITION Cloning vector pCDNA3ZEO DNA.
ACCESSION X90639
VERSION X90639.1 GI:949972
KEYWORDS cloning vector; expression vector; multiple cloning site; Plasmid.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Peters,H., Hunthausen,T., Kroenke,M. and Marget,M.
TITLE A new small sized high-level eukaryotic expression vector
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3986)
AUTHORS Peters,H.
TITLE Direct Submission
JOURNAL Submitted (07-AUG-1995) H. Peters, Inst. f. Immunologie,
Michaelstr.5, D- 24105 Kiel, FRG
COMMENT Related sequences: M21295 and K03104.
FEATURES
source
1. 3986
/organism="synthetic construct"
/mol_type="other DNA"
/db_xref="taxon:32630"
/plasmid="pCDNA3ZEO"
misc_feature 1..2125
/note="Cloning vector (pCDNA3) (Invitrogen)"
misc_feature 889..994
/note="multiple cloning site (MCS)"
misc_feature 2126..2796
/note="cloning vector (PzeoSV) (Invitrogen)"
misc_feature 2797..3986
/note="cloning vector (pCDNA3)"

ORIGIN
Query Match 100.0%; Score 100; DB 12; Length 3986;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
|
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
|
Qy 61 CCGCATAGTTAAGCAGTATCTGCTCCCTGCTTGTGTGTT 100
|
Db 61 CCGCATAGTTAAGCAGTATCTGCTCCCTGCTTGTGTGTT 100
|

RESULT 5
AR098191
LOCUS 4026 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 19 from patent US 6074850.
ACCESSION AR098191
VERSION AR098191.1 GI:12807448
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 4026)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion polypeptides
JOURNAL Patent: US 6074850-A 19 13-JUN-2000;
FEATURES
source
1. 4026
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4026;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
|
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
|
Qy 61 CCGCATAGTTAAGCAGTATCTGCTCCCTGCTTGTGTGTT 100
|
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
|
```

```

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100
    |||||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100

RESULT 6
AR207833
LOCUS AR207833 4026 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 19 from patent US 6379927.
ACCESSION AR207833
VERSION AR207833.1 GI:21507689
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 4026)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion proteins
JOURNAL Patent: US 6379927-A 19 30-APR-2002;
FEATURES
    Location/Qualifiers
        source
            1. 4026
                /organism="unknown"
                /mol_type="unassigned DNA"
ORIGIN
    Query Match 100.0%; Score 100; DB 6; Length 4026;
    Best Local Similarity 100.0%; Pred. No. 9.3e-24;
    Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
    |||||||
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60

RESULT 7
AR207833
LOCUS AR207833 4026 bp DNA linear PAT 31-JAN-2002
DEFINITION Tissue specific expression of retinoblastoma protein.
ACCESSION BD009730
VERSION BD009730.1 GI:18638103
KEYWORDS JP 2001503638-A/4.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 4026)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Tissue specific expression of retinoblastoma protein
JOURNAL Patent: JP 2001503638-A 4 21-MAR-2001;
COMMENT CANJI INC
OS Unidentified
PN JP 2001503638-A/4
PD 21-MAR-2001
PF 13-NOV-1997 JP 1998522958
PR 15-NOV-1996 US 08/751517,14-FEB-1997 US 08/801092 PI
DOUGLAS ANTELMAN, RICHARD J GREGORY, KENNETH N WILLS PC
C07H21/04, C07K5/00, A61K38/00, A61K35/12
CC Strandedness: Single;
CC Topology: Linear;
FH Key
FT source
    Location/Qualifiers
        1. 4026
            /organism="Unidentified"
            /mol_type="genomic DNA"
            /db_xref="taxon:32644"
FEATURES
    source
        1. 4026
            Location/Qualifiers
                /organism="unidentified"
                /mol_type="genomic DNA"
                /db_xref="taxon:32644"
ORIGIN
    Query Match 100.0%; Score 100; DB 6; Length 4026;
    Best Local Similarity 100.0%; Pred. No. 9.3e-24;
    Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
    |||||||
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100
    |||||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100

RESULT 8
AR207833
LOCUS AR207833 4249 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 33 from patent US 6074850.
ACCESSION AR098192
VERSION AR098192.1 GI:12807449
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 4249)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion polypeptides
JOURNAL Patent: US 6074850-A 33 13-JUN-2000;
FEATURES
    Location/Qualifiers
        source
            1. 4249
                /organism="unknown"
                /mol_type="unassigned DNA"
ORIGIN
    Query Match 100.0%; Score 100; DB 6; Length 4249;
    Best Local Similarity 100.0%; Pred. No. 9.3e-24;
    Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
    |||||||
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100
    |||||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100

RESULT 9
AR207834
LOCUS AR207834 4249 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 33 from patent US 6379927.
ACCESSION AR207834
VERSION AR207834.1 GI:21507691
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 4249)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion proteins
JOURNAL Patent: US 6379927-A 33 30-APR-2002;
FEATURES
    Location/Qualifiers
        source
            1. 4249
                /organism="unknown"
                /mol_type="unassigned DNA"
ORIGIN
    Query Match 100.0%; Score 100; DB 6; Length 4249;
    Best Local Similarity 100.0%; Pred. No. 9.3e-24;
    Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
    |||||||
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100
    |||||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTGTGTGTT 100

RESULT 10
AR207834
LOCUS AR207834 4249 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 33 from patent US 6379927.
ACCESSION AR207834
VERSION AR207834.1 GI:21507691
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 4249)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Retinoblastoma fusion proteins
JOURNAL Patent: US 6379927-A 33 30-APR-2002;
FEATURES
    Location/Qualifiers
        source
            1. 4249
                /organism="unknown"
                /mol_type="unassigned DNA"
ORIGIN
    Query Match 100.0%; Score 100; DB 6; Length 4249;
    Best Local Similarity 100.0%; Pred. No. 9.3e-24;
    Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
    |||||||
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60

```

```
Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

RESULT 10
BD009731
LOCUS BD009731 4249 bp DNA linear PAT 31-JAN-2002
DEFINITION Tissue specific expression of retinoblastoma protein.
ACCESSION BD009731
VERSION BD009731.1 GI:18638104
KEYWORDS JP 2001503638-A/5.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 4249)
AUTHORS Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE Tissue specific expression of retinoblastoma protein
JOURNAL Patent: JP 2001503638-A 5 21-MAR-2001;
CANJ I INC
COMMENT OS Unidentified
PN JP 2001503638-A/5
PD 21-MAR-2001
PF 13-NOV-1997 JP 1998522958
PR 15-NOV-1996 US 08/751517,14-FEB-1997 US 08/801092 PI
DOUGLAS ANTELMAN,RICHARD J GREGORY,KENNETH N WILLS PC
C07H21/04,C07K5/00,A61K38/00,A61K35/12
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1..4249 /organism='Unidentified'.
FEATURES
source
1..4249
Location/Qualifiers
/organism='unidentified'
/mol_type='genomic DNA'
/db_xref='taxon:32644'

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4249;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCGATCCCTATGGTGCAGTCTCAGTACAAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCGATCCCTATGGTGCAGTCTCAGTACAAATCTGCTCTGATG 60
Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

RESULT 11
A38214
LOCUS A38214 4341 bp DNA linear PAT 05-MAR-1997
DEFINITION Sequence 58 from Patent WO9408008.
ACCESSION A38214
VERSION A38214.1 GI:2294819
KEYWORDS
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 4341)
AUTHORS Hawkins,R.E., Russell,S.J., Stevenson,F.K. and Winter,G.P.
TITLE IMPROVEMENTS IN OR RELATING TO IMMUNE RESPONSE MODIFICATION
JOURNAL Patent: WO 9408008-A 58 14-APR-1994;
MEDICAL RES COUNCIL (GB)
COMMENT Other publication CA 2145064 940414
Other publication AU 4832493 940426
Other publication JP 8501699T 960227.
FEATURES
source
1..4341
Location/Qualifiers
/organism='unidentified'
/mol_type='genomic DNA'
/db_xref='taxon:32644'

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4249;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCGATCCCTATGGTGCAGTCTCAGTACAAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCGATCCCTATGGTGCAGTCTCAGTACAAATCTGCTCTGATG 60
Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

RESULT 12
AX286570
LOCUS AX286570 4341 bp DNA linear PAT 21-NOV-2001
DEFINITION Sequence 1 from Patent WO0179510.
ACCESSION AX286570
VERSION AX286570.1 GI:17048664
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Rice,J.H. and Stevenson,F.M.
TITLE Materials and methods relating to immune responses to fusion
JOURNAL Patent: WO 0179510-A 1 25-OCT-2001;
Cancer Research Ventures Limited (GB)
FEATURES
source
1..4341
Location/Qualifiers
/organism='synthetic construct'
/mol_type='unassigned DNA'
/db_xref='taxon:32630'
/notes='Vector pVAC1'

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4341;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCGATCCCTATGGTGCAGTCTCAGTACAAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCGATCCCTATGGTGCAGTCTCAGTACAAATCTGCTCTGATG 60
Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

RESULT 13
AX060344
LOCUS AX060344 4597 bp DNA linear PAT 22-JAN-2001
DEFINITION Sequence 3 from Patent WO0078358.
ACCESSION AX060344
VERSION AX060344.1 GI:12405832
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Chen,W.
TITLE Hyaluronic acid microspheres for sustained gene transfer
JOURNAL Patent: WO 0078358-A 3 28-DEC-2000;
The Collaborative Group, Ltd. (US)
FEATURES
source
1..4597
Location/Qualifiers
/organism='synthetic construct'
```

```
ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4341;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCGATCCCTATGGTGCAGTCTCAGTACAAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCGATCCCTATGGTGCAGTCTCAGTACAAATCTGCTCTGATG 60
Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

RESULT 12
AX286570
LOCUS AX286570 4341 bp DNA linear PAT 21-NOV-2001
DEFINITION Sequence 1 from Patent WO0179510.
ACCESSION AX286570
VERSION AX286570.1 GI:17048664
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Rice,J.H. and Stevenson,F.M.
TITLE Materials and methods relating to immune responses to fusion
JOURNAL Patent: WO 0179510-A 1 25-OCT-2001;
Cancer Research Ventures Limited (GB)
FEATURES
source
1..4341
Location/Qualifiers
/organism='synthetic construct'
/mol_type='unassigned DNA'
/db_xref='taxon:32630'
/notes='Vector pVAC1'

ORIGIN
Query Match 100.0%; Score 100; DB 6; Length 4341;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCGATCCCTATGGTGCAGTCTCAGTACAAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCGATCCCTATGGTGCAGTCTCAGTACAAATCTGCTCTGATG 60
Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

RESULT 13
AX060344
LOCUS AX060344 4597 bp DNA linear PAT 22-JAN-2001
DEFINITION Sequence 3 from Patent WO0078358.
ACCESSION AX060344
VERSION AX060344.1 GI:12405832
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Chen,W.
TITLE Hyaluronic acid microspheres for sustained gene transfer
JOURNAL Patent: WO 0078358-A 3 28-DEC-2000;
The Collaborative Group, Ltd. (US)
FEATURES
source
1..4597
Location/Qualifiers
/organism='synthetic construct'
```

```

/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="pCDNA3.1/GS vector by Invitrogen Corporation"

ORIGIN
Query Match      100.0%; Score 100; DB 6; Length 4597;
Best Local Similarity 100.0%; Pred. No. 9.2e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTCATG 60
    |||||||
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTCATG 60

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
    |||||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

RESULT 14
AX133940
LOCUS      AX133940      4840 bp      DNA      linear      PAT 15-MAY-2001
DEFINITION Sequence 1 from Patent WO0119853.
ACCESSION AX133940
VERSION    AX133940.1 GI:14139881
KEYWORDS   .
SOURCE     synthetic construct
           synthetic construct
           other sequences; artificial sequences.
REFERENCE  1
AUTHORS    Hollander,A.P., Barker,M.D. and Kafienah,W.C.
TITLE      Cell transfection
JOURNAL    Patent: WO 0119853-A 1 22-MAR-2001;
           THE UNIVERSITY OF SHEFFIELD (GB)
FEATURES   Location/Qualifiers
            source
              1..4840
              /organism="synthetic construct"
              /mol_type="unassigned DNA"
              /db_xref="taxon:32630"
              /note="This sequence is artificial and is based on well
              established commercially available vectors that are cited
              with their vendor within the patent application"

ORIGIN
Query Match      100.0%; Score 100; DB 6; Length 4840;
Best Local Similarity 100.0%; Pred. No. 9.2e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTCATG 60
    |||||||
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTCATG 60

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
    |||||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

RESULT 15
BD238492
LOCUS      BD238492      5053 bp      DNA      linear      PAT 17-JUL-2003
DEFINITION Expression vectors for stimulating an immune response and methods
           of using the same.
ACCESSION BD238492
VERSION    BD238492.1 GI:33048262
KEYWORDS   JP 2002520000-A/18.
SOURCE     synthetic construct
           synthetic construct
           other sequences; artificial sequences.
REFERENCE  1 (bases 1 to 5053)
AUTHORS    Fikes,J.D., Hermanson,G.G., Sette,A., Ishioka,G.Y., Livingston,B.
           and Chesnut,K.W.
TITLE      Expression vectors for stimulating an immune response and methods
           of using the same
JOURNAL    Patent: JP 2002520000-A 18 09-JUL-2002;

```

```

EPIMMUNE INC
OS Artificial Sequence
PN JP 2002520000-A/18
PD 09-JUL-2002
PF 13-MAY-1999 JP 2000548449
PR 13-MAY-1998 US 09/078904, 15-MAY-1998 US 60/085751 PI
JOHN D FIKES,GARY G HERMANSON,ALESSANDRO
SETTE,GLENN Y ISHIOKA,
PI BRIAN LIVINGSTON,ROBERT W CHESNUT
PC C12N15/09,A61K31/711,A61K39/12,A61K39/21,A61K39/29,
A61P31/14,
PC A61P31/16,A61P31/20,A61P37/02,C12N15/00
CC vector pEF2
FH Key Location/Qualifiers
FT source 1..5053
FT /organism="Artificial Sequence".

FEATURES
Source
  Location/Qualifiers
    1..5053
    /organism="synthetic construct"
    /mol_type="genomic DNA"
    /db_xref="taxon:32630"

ORIGIN
Query Match      100.0%; Score 100; DB 6; Length 5053;
Best Local Similarity 100.0%; Pred. No. 9.2e-24;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTCATG 60
    |||||||
Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTCATG 60

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
    |||||||
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

Search completed: July 14, 2005, 14:03:20
Job time : 752.127 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:35:42 ; Search time 140.988 Seconds
(without alignments)
4198.742 Million cell updates/sec

Title: US-09-482-682-8_COPY_1_100
Perfect score: 100
Sequence: 1 gacggatcggagatctccc.....ctgtccctctgtgtgtt 100

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100	100.0	1506	12	Adm41035 Fungus nu
2	100	100.0	1600	12	Adh11349 Vertebrat
3	100	100.0	1782	12	Adm41037 Cytomegal
4	100	100.0	2241	12	Adm41034 Human nuc
5	100	100.0	2294	12	Adm41036 Cytomegal
6	100	100.0	3853	2	Aav40006 Plasmid p
7	100	100.0	4026	2	Aav40007 Plasmid p
8	100	100.0	4249	2	Aav63466 Plasmid p
9	100	100.0	4341	2	Aag62391 Vector pv
10	100	100.0	4341	6	Aas17704 Vector pv
11	100	100.0	4341	6	Abs83143 Plasmid p
12	100	100.0	4597	4	Aaf24901 Nucleotid
13	100	100.0	4639	6	Aad39652 Human sma
14	100	100.0	4840	4	Aaf83146 Complete
15	100	100.0	5015	10	Adb33528 Expressio
16	100	100.0	5053	3	Aaz38633 pep2 expr
17	100	100.0	5070	4	Aas12839 DNA sequ
18	100	100.0	5082	2	Adh11417 Plasmid p
19	100	100.0	5162	10	Adf10526 Plasmid p
20	100	100.0	5162	10	Acc44637 Murine rd

21	100	100.0	5172	13	Ads75099 Plasmid p
22	100	100.0	5192	10	Acc44692 Plasmid p
23	100	100.0	5271	10	Abv77540 Plasmid p
24	100	100.0	5283	10	Abv77538 Plasmid p
25	100	100.0	5293	10	Abv77549 Plasmid p
26	100	100.0	5302	12	Adi34681 Nucleotid
27	100	100.0	5304	10	Abv77539 Plasmid p
28	100	100.0	5425	2	Adh11233 Vertebrat
29	100	100.0	5431	6	Abn86685 Plasmid v
30	100	100.0	5431	10	Ade21866 Plasmid v
31	100	100.0	5431	12	Ado05277 pcDNA3 pl
32	100	100.0	5432	3	Aaz89476 Transgeni
33	100	100.0	5446	6	Aas18619 Renilla l
34	100	100.0	5446	6	AbL53540 Vector pc
35	100	100.0	5446	12	Adn36314 Plasmid p
36	100	100.0	5458	6	AbL58494 Recombina
37	100	100.0	5458	6	AbL58493 Recombina
38	100	100.0	5543	6	AbK88868 Topoisome
39	100	100.0	5543	12	AdE83791 Plasmid p
40	100	100.0	5543	12	Ado06720 Recombina
41	100	100.0	5614	6	AbL58489 Recombina
42	100	100.0	5614	6	AbL58490 Recombina
43	100	100.0	5618	2	Aaq88310 Plasmid p
44	100	100.0	5651	5	Aai66195 Human FSH
45	100	100.0	5651	6	AbK40237 DNA encod

ALIGNMENTS

RESULT 1

- ADm41035
- ID ADm41035 standard; DNA; 1506 BP.
- AC ADm41035;
- XX
- DT 17-JUN-2004 (first entry)
- DE Fungus nucleotide sequence SEQ ID NO:3.
- XX
- KW engrafting foreign replacement cell; implanting foreign replacement cell; growth; differentiation; drug development; vaccine development; tissue transplantation; human disease study; fungus; gene; ds.
- XX
- OS Unidentified.
- XX
- FN WO2004027029-A2.
- XX
- PD 01-APR-2004.
- XX
- PF 17-SEP-2003; 2003WO-US029251.
- XX
- PR 19-SEP-2002; 2002US-0411790P.
- XX
- PA (XIME-) XIMEREX INC.
- XX
- PI Beschorner WE, Sosa CE, Thompson SC;
- XX
- DR WPI; 2004-295402/27.
- XX
- PT Engrafting foreign replacement cells within a fetal non-human mammal, useful in producing chimeric mammals, comprises selectively destroying native cells in a tissue of a fetal non-human mammal host.
- XX
- XX Disclosure; SEQ ID NO 3; 48pp; English.
- XX
- CC The present invention describes a method for engrafting foreign replacement cells within a foetal non-human mammal, which comprises selectively destroying native cells in a tissue of a foetal non-human mammal host, where the number of maternal cells of the same tissue is not substantially reduced, and implanting foreign replacement cells in the tissue of the fetal non-human mammal host, where the foreign replacement cells replace destroyed cells of the tissue. The method is useful for

CC facilitating growth and differentiation of foreign cells within a
 CC mammalian host, and for producing chimeric mammals that can be used to
 CC develop new drugs and vaccine, factors, drugs and tissues for
 CC transplantation, also useful to study human diseases. The present
 CC sequence represents a nucleotide sequence given in the Sequence Listing
 CC of the present invention but not mentioned further within the
 CC specification.

XX
 SQ Sequence 1506 BP; 454 A; 277 C; 361 G; 414 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 12; Length 1506;
 Best Local Similarity 100.0%; Pred. No. 4e-26;
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
 |
 Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
 |
 Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
 |
 Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
 |

RESULT 2
 ADH11349
 ID ADH11349 standard; DNA; 1600 BP.
 XX
 AC ADH11349;
 XX
 DT 11-MAR-2004 (first entry)
 XX
 DE Vertebrate UNC-53 protein homologue related nucleotide sequence.

XX
 KW UNC-53 vertebrate protein homologue; UNC-53; Caenorhabditis elegans;
 KW cell shape regulator; cell motility regulator; cell migration;
 KW cell behaviour regulator; phenotype; signal transduction pathway;
 KW signal transducing protein; signal integrator protein;
 KW neuronal regeneration; revascularisation; wound healing;
 KW chronic neurodegenerative disease; acute traumatic injury;
 KW fibrotic disease; gene; ds.

XX
 OS Unidentified.
 XX
 PN WO9824810-A2.
 XX
 PD 11-JUN-1998.
 XX
 XX
 PF 03-DEC-1997; 97WO-EP006956.
 XX
 PR 04-DEC-1996; 96GB-00025283.
 XX
 XX (JANC) JANSSEN PHARM NV.
 XX
 XX Platteeuw CJ, Buesa Arjol CM, Deraeymaeker M, Verbasselt P;
 XX Pujol NJR, Maertens LJS, Luyten W, Geerts H, Vandekerckhove JS;
 XX Geysen J, Bogaert TAOR;
 XX
 XX WPI; 1998-362411/31.
 XX P-PSDB; ADH11350.
 XX
 XX Vertebrate homologues of C. elegans UNC-53 protein - useful for, e.g.
 XX promoting neuronal regeneration, treating chronic neuro-degenerative
 XX diseases or acute traumatic injuries.
 XX
 XX Disclosure; Page 410-411; 479pp; English.

XX
 CC The present invention describes a vertebrate protein homologue of an UNC-
 CC 53 protein or Caenorhabditis elegans or a functional equivalent,
 CC derivative or bioprecursor of UNC-53. Also described: (1) a cDNA sequence
 CC encoding a vertebrate homologue of the C. elegans UNC-53 protein; (2) a
 CC nucleic acid which hybridises to the cDNA of (1); (3) vector comprising
 CC the cDNA as in (1); (4) a host cell containing the vector as in (3); (5)
 CC a transgenic cell, tissue or animal comprising the vector as in (3); (6)

CC a compound identified as an enhancer or inhibitor of the regulation of
 CC cell shape, motility, or the direction of cell migration for use as a
 CC therapeutic; (7) a method for determination of whether a protein is an
 CC inhibitor or enhancer of regulation of cell behaviour, growth, shape or
 CC motility or the direction of migration by contacting a host cell
 CC expressing a homologue of UNC-53 and determining a change of phenotype;
 CC (8) a method for identification of vertebrate homologues of C. elegans
 CC unc-53 gene by hybridising a probe of 15-50 bp of the unc-53 sequence to
 CC a DNA library; and (9) a method for identification of a protein which is
 CC active in the signal transduction pathway of a cell of which a vertebrate
 CC homologue of UNC-53 is a component comprising: (i) contacting an extract
 CC of a cell with an antibody to the UNC-53 homologue; (ii) identifying an
 CC antibody/homologue complex; and (iii) analysing such a complex to
 CC identify any non-antibody protein bound to the complex. UNC-53 is a
 CC signal transducing or signal integrator protein involved in controlling
 CC directionality of cell migration and cell shape in C. elegans. Vertebrate
 CC homologues of UNC-53 can be used to promote neuronal regeneration,
 CC revascularisation or wound healing, to treat chronic neurodegenerative
 CC diseases or acute traumatic injuries or fibrotic diseases. The present
 CC sequence is used in the exemplification of the present invention.

XX
 SQ Sequence 1600 BP; 397 A; 409 C; 398 G; 396 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 2; Length 1600;
 Best Local Similarity 100.0%; Pred. No. 4.1e-26;
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
 |
 Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
 |
 Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
 |
 Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
 |

RESULT 3
 ADH41037
 ID ADH41037 standard; DNA; 1782 BP.
 XX
 AC ADH41037;
 XX
 DT 17-JUN-2004 (first entry)
 XX
 DE Cytomegalovirus nucleotide sequence SEQ ID NO:5.
 XX
 KW engrafting foreign replacement cell; implanting foreign replacement cell;
 KW growth; differentiation; drug development; vaccine development;
 KW tissue transplantation; human disease study; cytomegalovirus; gene; ds.

XX
 OS Cytomegalovirus.
 XX
 PN WO2004027029-A2.
 XX
 PD 01-APR-2004.
 XX
 PF 17-SEP-2003; 2003WO-US029251.
 XX
 PR 19-SEP-2002; 2002US-0411790P.
 XX
 XX (XIME-) XIMEREX INC.
 XX
 XX Beschoner WE, Sosa CE, Thompson SC;
 XX WPI; 2004-295402/27.
 XX
 XX Engrafting foreign replacement cells within a fetal non-human mammal,
 XX useful in producing chimeric mammals, comprises selectively destroying
 XX native cells in a tissue of a fetal non-human mammal host.

XX
 PS Disclosure; SEQ ID NO 5; 48pp; English.
 XX
 CC The present invention describes a method for engrafting foreign

CC replacement cells within a foetal non-human mammal, which comprises
CC selectively destroying native cells in a tissue of a foetal non-human
CC mammal host, where the number of maternal cells of the same tissue is not
CC substantially reduced, and implanting foreign replacement cells in the
CC tissue of the foetal non-human mammal host, where the foreign replacement
CC cells replace destroyed cells of the tissue. The method is useful for
CC facilitating growth and differentiation of foreign cells within a
CC mammalian host, and for producing chimeric mammals that can be used to
CC develop new drugs and vaccine, factors, drugs and tissues for
CC transplantation, also useful to study human diseases. The present
CC sequence represents a nucleotide sequence given in the Sequence Listing
CC of the present invention but not mentioned further within the
CC specification.

XX SQ Sequence 1782 BP; 458 A; 399 C; 451 G; 474 T; 0 U; 0 Other;
Query Match 100.0%; Score 100; DB 12; Length 1782;
Best Local Similarity 100.0%; Pred. No. 4.2e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTTGATG 60
DB 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTTGATG 60
OY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100
DB 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100

RESULT 4

ADM41034
ID ADM41034 standard; DNA; 2241 BP.

XX AC ADM41034;

XX DT 17-JUN-2004 (first entry)

XX XX Human nucleotide sequence SEQ ID NO:2.

XX engrafting foreign replacement cell; implanting foreign replacement cell;
XX growth; differentiation; drug development; vaccine development;
XX tissue transplantation; human disease study; human; gene; ds.

XX OS Homo sapiens.

XX PN WO2004027029-A2.

XX PD 01-APR-2004.

XX PF 17-SEP-2003; 2003WO-US029251.

XX PR 19-SEP-2002; 2002US-0411790P.

XX PA (XIME-) XIMEREX INC.

XX PI Beschornor WE, Sosa CE, Thompson SC;

XX DR WPI; 2004-295402/27.

XX Engrafting foreign replacement cells within a foetal non-human mammal,
XX useful in producing chimeric mammals, comprises selectively destroying
XX native cells in a tissue of a foetal non-human mammal host.

XX PS Disclosure; SEQ ID NO 2; 48pp; English.

XX The present invention describes a method for engrafting foreign
XX replacement cells within a foetal non-human mammal, which comprises
XX selectively destroying native cells in a tissue of a foetal non-human
XX mammal host, where the number of maternal cells of the same tissue is not
XX substantially reduced, and implanting foreign replacement cells in the
XX tissue of the foetal non-human mammal host, where the foreign replacement
XX cells replace destroyed cells of the tissue. The method is useful for
XX facilitating growth and differentiation of foreign cells within a

CC mammalian host, and for producing chimeric mammals that can be used to
CC develop new drugs and vaccine, factors, drugs and tissues for
CC transplantation, also useful to study human diseases. The present
CC sequence represents a nucleotide sequence given in the Sequence Listing
CC of the present invention but not mentioned further within the
CC specification.

XX SQ Sequence 2241 BP; 498 A; 630 C; 609 G; 504 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 12; Length 2241;
Best Local Similarity 100.0%; Pred. No. 4.5e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTTGATG 60
DB 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTTGATG 60

OY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100

DB 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100

RESULT 5

ADM41036
ID ADM41036 standard; DNA; 2294 BP.

XX AC ADM41036;

XX DT 17-JUN-2004 (first entry)

XX DE Cytomegalovirus nucleotide sequence SEQ ID NO:4.

XX engrafting foreign replacement cell; implanting foreign replacement cell;
XX growth; differentiation; drug development; vaccine development;
XX tissue transplantation; human disease study; cytomegalovirus; gene; ds.

XX OS Cytomegalovirus.

XX PN WO2004027029-A2.

XX PD 01-APR-2004.

XX PF 17-SEP-2003; 2003WO-US029251.

XX PR 19-SEP-2002; 2002US-0411790P.

XX PA (XIME-) XIMEREX INC.

XX PI Beschornor WE, Sosa CE, Thompson SC;

XX DR WPI; 2004-295402/27.

XX Engrafting foreign replacement cells within a foetal non-human mammal,
XX useful in producing chimeric mammals, comprises selectively destroying
XX native cells in a tissue of a foetal non-human mammal host.

XX PS Disclosure; SEQ ID NO 4; 48pp; English.

XX The present invention describes a method for engrafting foreign
XX replacement cells within a foetal non-human mammal, which comprises
XX selectively destroying native cells in a tissue of a foetal non-human
XX mammal host, where the number of maternal cells of the same tissue is not
XX substantially reduced, and implanting foreign replacement cells in the
XX tissue of the foetal non-human mammal host, where the foreign replacement
XX cells replace destroyed cells of the tissue. The method is useful for
XX facilitating growth and differentiation of foreign cells within a
XX mammalian host, and for producing chimeric mammals that can be used to
XX develop new drugs and vaccine, factors, drugs and tissues for
XX transplantation, also useful to study human diseases. The present
XX sequence represents a nucleotide sequence given in the Sequence Listing
XX of the present invention but not mentioned further within the
XX specification.

```
SQ Sequence 2294 BP; 466 A; 680 C; 641 G; 507 T; 0 U; 0 Other;

Query Match      100.0%; Score 100; DB 12; Length 2294;
Best Local Similarity 100.0%; Pred. No. 4.5e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100

RESULT 6
AAV40006
ID AAV40006 standard; DNA; 3853 BP.
XX
AC AAV40006;
XX
XX
DT 27-AUG-2003 (revised)
DT 15-FEB-1999 (first entry)
XX
DE Plasmid pCTM.
XX
KW E2F; transcription factor; human; retinoblastoma protein RB;
KW bladder cancer; restenosis; angioplasty; diabetic retinopathy;
KW thyroid hyperplasia; Grave's disease; psoriasis;
KW benign prostatic hypertrophy; Li-Fraumeni syndrome;
KW peripheral vascular disease; therapy; plasmid pCTM; ss.
XX
OS Human cytomegalovirus.
OS mastadenovirus.
OS unidentified bacteriophage; T7.
OS unidentified bacteriophage; SP6.
OS Macaca mulatta; polyoma virus.
OS Bos taurus.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT promoter 209..864
FT /*tag= a
FT /*note= "CMV promoter"
FT misc_feature 907..1131
FT /*tag= b
FT /*function= "tripartite leader sequence"
FT promoter 1132..1149
FT /*tag= c
FT /*note= "SP6 promoter"
FT misc_feature 1679..3853
FT /*tag= d
FT /*note= "pUC19 backbone H3 to AatII"
FT CDS complement(2857..3717)
FT /*tag= e
FT /*note= "AMP-ORF"
XX
XX
PN WO9821228-A1.
XX
XX
PD 22-MAY-1998.
XX
PF 13-NOV-1997; 97WO-US021821.
XX
XX
PR 15-NOV-1996; 96US-00751517.
PR 14-FEB-1997; 97US-00801092.
XX
XX
PA (CANJ-) CANJI INC.
XX
XX
PI Antelman D, Gregory RJ, Wills KN;
XX
XX
DR WPI; 1998-297858/26.
XX
XX
PT New fusion polypeptide of, e.g. transcription factor - used to treat,
```

```
PT e.g. hyper-proliferative disease such as cancer and restenosis.
XX
XX
XX Example 1; Fig 4; 91pp; English.
XX
CC This is the nucleotide sequence of pCTM, a plasmid which contains a CMV
CC promoter, a tripartite adenovirus leader flanked by T7 and SP6 promoters,
CC and a multiple cloning site with a bovine growth hormone polyA site and
CC downstream SV40 polyA site. It has been used as a vector for the
CC expression of fusion proteins of the invention that comprise
CC retinoblastoma protein (BP, see AAW62465) and E2F transcription factor
CC (see AAW62464). Such fusion proteins, particularly expressed from gene
CC therapy vectors, are used to treat hyperproliferative conditions,
CC specifically cancer (particularly of the bladder) or restenosis. They are
CC more effective in repressing transcription of the E2F promoter than RB
CC alone and cause cell-cycle arrest in a variety of cells. (Updated on 27-
CC AUG-2003 to correct OS field.)
XX
SQ Sequence 3853 BP; 936 A; 986 C; 942 G; 989 T; 0 U; 0 Other;

Query Match      100.0%; Score 100; DB 2; Length 3853;
Best Local Similarity 100.0%; Pred. No. 5.2e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60

Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTT 100

RESULT 7
AAV40007
ID AAV40007 standard; DNA; 4026 BP.
XX
AC AAV40007;
XX
XX
DT 27-AUG-2003 (revised)
DT 15-FEB-1999 (first entry)
XX
DE Plasmid pCTMI.
XX
KW E2F; transcription factor; human; retinoblastoma protein RB;
KW bladder cancer; restenosis; angioplasty; diabetic retinopathy;
KW thyroid hyperplasia; Grave's disease; psoriasis;
KW benign prostatic hypertrophy; Li-Fraumeni syndrome;
KW peripheral vascular disease; therapy; plasmid pCTMI; ss.
XX
OS Human cytomegalovirus.
OS mastadenovirus.
OS unidentified bacteriophage; T7.
OS unidentified bacteriophage; SP6.
OS Macaca mulatta; polyoma virus.
OS Bos taurus.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT promoter 209..864
FT /*tag= a
FT /*note= "CMV promoter"
FT misc_feature 907..1074
FT /*tag= b
FT /*function= "tripartite leader sequence"
FT intron 1075..1253
FT /*tag= c
FT /*note= "hybrid SV40 late intron"
FT promoter 1305..1322
FT /*tag= d
FT /*note= "SP6 promoter"
FT misc_feature 1851..4026
FT /*tag= e
FT /*note= "pUC19 backbone H3 to AatII"
```

```

FT CDS complement (3032..3890)
FT /*tag= f
FT /*note= "AMP-ORF"
PN WO9821228-A1.
PD 22-MAY-1998.
XX 13-NOV-1997; 97WO-US021821.
XX 15-NOV-1996; 96US-00751517.
XX 14-FEB-1997; 97US-00801092.
XX (CANJ-) CANJI INC.
XX Antelman D, Gregory RJ, Wills KN;
XX WPI; 1998-297858/26.
XX New fusion polypeptide of, e.g. transcription factor - used to treat,
XX e.g. hyper-proliferative disease such as cancer and restenosis.
XX Example 1; Fig 6; 91pp; English.
XX This is the nucleotide sequence of pCTMI, a plasmid that was constructed
XX from pCTM (see AAV40006) by digesting pCTM with XhoI and NotI and
XX subcloning a 180 bp intron XhoI-NotI fragment from a pCMV-beta-gal
XX vector. Plasmid pCTMI has been used as a vector for the expression of
XX fusion proteins of the invention that comprise retinoblastoma protein
XX (BP, see AAW62465) and E2F transcription factor (see AAW62464). Such
XX fusion proteins, particularly expressed from gene therapy vectors, are
XX used to treat hyperproliferative conditions, specifically cancer
XX (particularly of the bladder) or restenosis. They are more effective in
XX repressing transcription of the E2F promoter than RB alone and cause cell
XX -cycle arrest in a variety of cells. (Updated on 27-AUG-2003 to correct
XX OS field.)
XX SQ Sequence 4026 BP; 978 A; 1021 C; 982 G; 1045 T; 0 U; 0 Other;

Query Match 100.0%; Score 100; DB 2; Length 4026;
Best Local Similarity 100.0%; Pred. No. 5.3e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACATCTGCTCTGATG 60
Db 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACATCTGCTCTGATG 60
Qy 61 CCGCATAGTTAAGCAGTATCTGCTCCCTGCTTGTGTGTT 100
Db 61 CCGCATAGTTAAGCAGTATCTGCTCCCTGCTTGTGTGTT 100

RESULT 8
AAV63466
ID AAV63466 standard; DNA; 4249 BP.
XX
XX AC AAV63466;
XX
XX 27-AUG-2003 (revised)
XX 15-FEB-1999 (first entry)
XX Plasmid pCTMIE.
XX E2F; transcription factor; human; retinoblastoma protein RB;
XX bladder cancer; restenosis; angioplasty; diabetic retinopathy;
XX thyroid hyperplasia; Grave's disease; psoriasis;
XX benign prostatic hypertrophy; Li-Fraumeni syndrome;
XX peripheral vascular disease; therapy; plasmid pCTMIE; ss.
XX Human cytomegalovirus.
OS mastadenovirus.
OS unidentified bacteriophage; T7.
OS unidentified bacteriophage; SP6.

```

```

OS Macaca mulatta; polyoma virus.
OS Bos taurus.
OS Chimeric.
XX Key Location/Qualifiers
XX promoter 209..864
XX /*tag= a
XX /*note= "CMV promoter"
XX misc_feature 907..1074
XX /*tag= b
XX /function= "tripartite leader sequence"
XX intron 1081..1145
XX /*tag= c
XX /*note= "hybrid SV40 late intron"
XX mRNA 1164..1366
XX /*tag= d
XX /*note= "early mRNA"
XX enhancer 1261..1332
XX /*tag= e
XX /*note= "72 bp tandem repeat enhancer"
XX enhancer 1333..1404
XX /*tag= f
XX /*note= "72 bp tandem repeat enhancer"
XX misc_binding 1366
XX /*tag= g
XX /*note= "T antigen binding site"
XX intron 1372..1478
XX /*tag= h
XX /*note= "hybrid SV40 late intron"
XX promoter 1530..1545
XX /*tag= i
XX /*note= "SP6 promoter"
XX misc_feature 2075..4249
XX /*tag= j
XX /*note= "pUC19 backbone H3 to AatII"
XX CDS complement(3255..4113)
XX /*tag= k
XX /*note= "AMP-ORF"
XX
XX WO9821228-A1.
XX
XX 22-MAY-1998.
XX
XX 13-NOV-1997; 97WO-US021821.
XX
XX 15-NOV-1996; 96US-00751517.
XX 14-FEB-1997; 97US-00801092.
XX (CANJ-) CANJI INC.
XX
XX Antelman D, Gregory RJ, Wills KN;
XX
XX WPI; 1998-297858/26.
XX
XX New fusion polypeptide of, e.g. transcription factor - used to treat,
XX e.g. hyper-proliferative disease such as cancer and restenosis.
XX Example 1; Fig 8; 91pp; English.
XX This is the nucleotide sequence of pCTMIE, a plasmid that was constructed
XX by amplifying the SV40 enhancer from SV40 viral DNA by PCR, digesting the
XX amplified product with BglII and inserting into BamHI-digested plasmid
XX pCMTI (see AAV40007). Plasmid pCTMIE has been used as a vector for the
XX expression of fusion proteins of the invention that comprise
XX retinoblastoma protein (BP, see AAW62465) and E2F transcription factor
XX (see AAW62464). Such fusion proteins, particularly expressed from gene
XX therapy vectors, are used to treat hyperproliferative conditions,
XX specifically cancer (particularly of the bladder) or restenosis. They are
XX more effective in repressing transcription of the E2F promoter than RB
XX alone and cause cell-cycle arrest in a variety of cells. (Updated on 27-
XX AUG-2003 to correct OS field.)
XX
XX Sequence 4249 BP; 1020 A; 1074 C; 1048 G; 1107 T; 0 U; 0 Other;

```

Query Match 100.0%; Score 100; DB 2; Length 4249;
 Best Local Similarity 100.0%; Pred. No. 5.4e-26;
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
 |||||
 Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
 |||||

Qy 61 CCGCATAGTTAAGCCAGTATCTGTCCTGCTTGTGTGTT 100
 |||||
 Db 61 CCGCATAGTTAAGCCAGTATCTGTCCTGCTTGTGTGTT 100
 |||||

RESULT 9
 AAQ62391
 ID AAQ62391 standard; DNA; 4341 BP.
 XX
 AC AAQ62391;
 XX
 DT 25-MAR-2003 (revised)
 DT 18-NOV-1994 (first entry)
 XX
 DE Vector pVAC1.
 XX
 KW Vector; pVAC1; pRC/RSV; leader sequence; termination signal;
 KW fusion protein; pSfi/Not.Tag1; pElB leader; human; immunoglobulin; VH1;
 KW single chain; Fv; murine antibody; retroviral; envelope; plasmid;
 KW vaccine; ss.
 XX
 OS Synthetic.
 XX

FH Key Location/Qualifiers
 FT misc_RNA complement(1. .775)
 FT /*tag= c
 FT /*note= "Claim 9"
 FT 606. .780
 FT /*tag= b
 FT /*note= "Claim 8"
 FT 606. .716
 FT /*tag= a
 FT /*note= "Claim 7"
 XX
 PN WO9408008-A1.
 XX
 PD 14-APR-1994.
 XX
 PF 04-OCT-1993; 93WO-GB002054.
 XX
 PR 02-OCT-1992; 92GB-00020808.
 XX
 PA (MEDI-) MEDICAL RES COUNCIL.
 XX
 PI Hawkins RE, Russell SJ, Stevenson FK, Winter GP;
 XX
 DR WPI; 1994-135575/16.
 XX

PT Modulating immune response to a disease marker - by administering a
 PT vector which expresses the disease marker to interact with the immune
 PT system.
 XX
 PS Claim 10; Fig 7; 77pp; English.
 XX

CC This sequence represents the vector pVAC1. This vector is based on the
 CC commercially available vector pRC/RSV. Leader sequences and termination
 CC signals were introduced into the vector to allow for production of fusion
 CC proteins. The vector, pSfi/Not.Tag1, was modified to replace the pElB
 CC leader with the human immunoglobulin VH1 leader sequence that permits the
 CC encoding of an SfiI cloning site without modification of the amino acid
 CC sequence. This fragment was then cloned as an EcoRI/Blunt-HindIII
 CC fragment into NotI/Blunt- HindIII cut vector pRC/RSV to give pVAC1. The
 CC single chain Fv for an individual patient can be inserted within the VH1
 CC leader sequence. This plasmid when encoding a single chain murine

CC antibody/retroviral envelope fusion protein can be used as a plasmid
 CC vaccine and it induces a strong humoral response to the antibody moiety
 CC in BALB/c mice. (Updated on 25-MAR-2003 to correct PN field.)
 XX

SQ Sequence 4341 BP; 1032 A; 1099 C; 1091 G; 1119 T; 0 U; 0 Other;
 Query Match 100.0%; Score 100; DB 2; Length 4341;
 Best Local Similarity 100.0%; Pred. No. 5.4e-26;
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
 |||||
 Db 1 GACGGATCGGAGATCTCCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
 |||||

Qy 61 CCGCATAGTTAAGCCAGTATCTGTCCTGCTTGTGTGTT 100
 |||||
 Db 61 CCGCATAGTTAAGCCAGTATCTGTCCTGCTTGTGTGTT 100
 |||||

RESULT 10
 AAS17704
 ID AAS17704 standard; DNA; 4341 BP.
 XX
 AC AAS17704;
 XX
 DT 12-MAR-2002 (first entry)
 XX
 DE Vector pVAC1 encoding a DNA vaccine.
 XX
 KW Cytostatic; vaccine; tetanus toxin; FrC; tumour; CTL; PCR primer; pVAC1;
 KW ds.
 XX
 OS Clostridium tetani.
 OS Homo sapiens.
 OS Synthetic.
 OS Cauliflower mosaic virus.
 XX
 PN WO200179510-A1.
 XX
 PD 25-OCT-2001.
 XX
 PF 17-APR-2001; 2001WO-GB001719.
 XX
 PR 17-APR-2000; 2000GB-00009470.
 XX
 PA (CANC-) CANCER RES VENTURES LTD.
 XX
 PI Rice J, Stevenson F;
 XX
 DR WPI; 2002-066370/09.
 XX

PT Nucleic acid construct, useful to immunize against various diseases
 PT including cancer, expresses the first domain of tetanus toxin FrC fused
 PT to a disease peptide antigen to provide a vaccine.
 XX
 PS Disclosure; Fig 4; 71pp; English.
 XX

CC The invention relates to a nucleic acid construct for delivery into
 CC living cells in vivo, to induce an immune response to a disease peptide
 CC antigen, where the construct directs expression of a fusion protein
 CC comprising the peptide antigen and the first domain of FrC. Also included
 CC are a nucleic acid vector comprising the above construct, a host cell
 CC comprising the above construct or vector and a method of producing a
 CC nucleic acid construct for inducing an immune response. The method
 CC comprises identifying a nucleic acid sequence encoding a disease peptide
 CC antigen comprising epitopes characteristic of the disease, cloning the
 CC nucleic acid sequence, introducing the cloned nucleic acid into a vector
 CC which allows the antigen to be expressed as a fusion with a first domain
 CC FrC from tetanus toxin, and optionally isolating the construct from the
 CC vector. The construct or vector is used as a vaccine to induce an immune
 CC response, particularly to tumour antigens. The present sequence is vector
 CC pVAC1 which encodes a vaccine of the invention
 XX

SQ Sequence 4341 BP; 1033 A; 1099 C; 1090 G; 1119 T; 0 U; 0 Other;
Query Match 100.0%; Score 100; DB 6; Length 4341;
Best Local Similarity 100.0%; Pred. No. 5.4e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
DB 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
DB 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

RESULT 11
ID ABN83143 standard; DNA; 4341 BP.
XX
AC ABN83143;
XX
DT 10-SEP-2002 (first entry)
XX
DE Plasmid pVAC1 complete sequence.
XX
KW Immune response; plant viral coat protein; pVAC1; cytostatic; virucide;
KW cancer; B cell malignancy; ds.
XX
OS Synthetic.
XX
FN WO200240513-A2.
XX
PD 23-MAY-2002.
XX
PF 20-NOV-2001; 2001WO-GB005142.
XX
PR 20-NOV-2000; 2000GB-00028319.
XX
PA (CANC-) CANCER RES VENTURES LTD.
XX
PI Savelyeva N, Stevenson F;
XX
DR WPI; 2002-500202/53.
XX
PT Nucleic acid construct for delivery into living cells as a vaccine,
PT useful for treating e.g. cancer, directs the expression of a fusion
PT protein comprising an antigen and an adjuvant sequence derived from a
PT plant viral coat protein.
XX
PS Example 3; Fig 7; 84pp; English.
XX
CC The invention relates to a novel nucleic acid construct for inducing an
CC immune response in vivo to an antigen, capable of directing the
CC expression of a fusion protein that comprises an antigen and an adjuvant
CC sequence derived from a plant viral coat protein. The construct of the
CC invention has cytostatic and virucide activity. The nucleic acid
CC construct is useful for inducing an immune response in a patient, for
CC vaccinating a patient against an infectious disease caused by an antigen
CC derived from a pathogen e.g. a virus, for treating a cancer patient or a
CC patient with a predisposition to cancer and for treating a patient having
CC a B cell malignancy, where the construct is encapsidated, and optionally,
CC a second nucleic acid sequence encoding a further immunomodulatory
CC polypeptide is administered to the patient. The construct is also useful
CC in medical treatment, and in the preparation of a vaccine for treating or
CC preventing a disease state associated with the antigen. The sequence
CC shows the complete sequence of vector pVAC1
XX

SQ Sequence 4341 BP; 1033 A; 1099 C; 1090 G; 1119 T; 0 U; 0 Other;
Query Match 100.0%; Score 100; DB 6; Length 4341;
Best Local Similarity 100.0%; Pred. No. 5.4e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
DB 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
DB 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

RESULT 12
ID AAF24901 standard; DNA; 4597 BP.
XX
AC AAF24901;
XX
DT 20-APR-2001 (first entry)
XX
DE Nucleotide sequence of the plasmid pCDNA3.1/GS.
XX
KW Microsphere; dihydrazide; hyaluronic acid; inflammatory response;
KW myocardial ischemia; cardiac angiogenesis; haemophilia;
KW vascular endothelial growth factor; VEGF; ss.
XX
OS Synthetic.
XX
FN WO200078358-A2.
XX
PD 28-DEC-2000.
XX
PF 19-JUN-2000; 2000WO-US016937.
XX
PR 18-JUN-1999; 99US-0140260P.
XX
PA (COLL-) COLLABORATIVE GROUP LTD.
XX
PI Chen W;
XX
DR WPI; 2001-071363/08.
XX
PT Hyaluronic acid micro spheres for use in gene therapy of myocardial
PT ischemia and hemophilia, comprising dihydrazide derivatized hyaluronic
PT acids crosslinked to nucleic acids.
XX
PS Example 1; Page 36-38; 38pp; English.
XX
CC The specification describes a microsphere comprising dihydrazide
CC derivatized hyaluronic acid crosslinked to a nucleic acid (NA). The
CC microspheres cause reduced inflammatory responses, and have increased
CC safety and biodegradability. The microspheres are useful for transfecting
CC a cell of a subject and for treating a subject having myocardial
CC ischemia, by increasing cardiac angiogenesis. They are also useful for
CC treating haemophilia. The present sequence represents the plasmid
CC pCDNA3.1/GS, into which is inserted a polynucleotide sequence which is
CC crosslinked to hyaluronic acid. The polynucleotide sequence encodes a
CC vascular endothelial growth factor (VEGF)
XX
SQ Sequence 4597 BP; 1062 A; 1214 C; 1206 G; 1115 T; 0 U; 0 Other;
Query Match 100.0%; Score 100; DB 4; Length 4597;
Best Local Similarity 100.0%; Pred. No. 5.5e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60
DB 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTGATG 60

QY 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
DB 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

RESULT 13
AAD39652

AD39652 standard; DNA; 4639 BP.
AD39652;
22-OCT-2002 (first entry)
Human small nuclear RNA (snRNA) DNA.
Human; recombinant vector; insertion cassette; small nuclear RNA; snRNA;
transgenic animal; ds.
Homo sapiens.
US2002058287-A1.
16-MAY-2002.
12-MAR-2001; 2001US-00804481.
10-MAR-2000; 2000US-0188304P.
(WHED) WHITEHEAD INST BIOMEDICAL RES.
Graaf DD, Lander ES;
WPI; 2002-499510/53.
New recombinant vector containing sequence for small nuclear RNA, useful
e.g. for identifying variant snRNA that suppresses expression of
transcription products.
Disclosure; Fig 1; 18pp; English.
The invention relates to a recombinant vector which comprises DNA,
consisting of an insertion cassette contained between at least two
insertion sites, that encodes a small nuclear (sn) RNA. The invention is
used to identify snRNA modifications that inhibit expression of
transcription products (and the identified snRNA are used to suppress
expression) for delivering antisense sequences to the nucleus and to
create transgenic animals. The present DNA sequence is human snRNA, U1
Sequence 4639 BP; 1067 A; 1198 C; 1243 G; 1131 T; 0 U; 0 Other;
Query Match 100.0%; Score 100; DB 6; Length 4639;
Best Local Similarity 100.0%; Pred. No. 5.5e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTCATG 60
Db 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTCATG 60
Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
RESULT 14
AAF83146
ID AAF83146 standard; DNA; 4840 BP.
AC AAF83146;
XX
DT 09-JUL-2001 (first entry)
DE Complete sequence of vector pIRES/BS.
XX
KW Blastocidin resistance; BS gene; gene therapy; tissue engineering;
KW cosmetic surgery; arthritis; joint replacement; skin graft; rhinoplasty;
KW pIRES/BS; ss.
XX
OS Synthetic.
XX
PN WO200119853-A2.

22-MAR-2001.
11-SEP-2000; 2000WO-GB003462.
11-SEP-1999; 99GB-00021418.
(UYSH-) UNIV SHEFFIELD.
Hollander AP, Barker MD, Kafienah W;
WPI; 2001-290354/30.
Novel nucleic acid molecule useful for therapeutic and cosmetic tissue
engineering, comprising at least a functional part of blastocidin
resistance gene linked through a recognition sequence, to a selected
gene.
Claim 11; Fig C; 44pp; English.
The invention provides a nucleic acid molecule (I) comprising at least
the functional part of blastocidin resistance (BS) gene, or its homolog,
linked through a recognition sequence to at least one selected gene. (I)
is useful in treatment comprising: (1) providing cells/tissues transfected
with (I); (2) surgical administration of the cells/tissues to the patient
to be treated; and optionally (3) monitoring the status of the cells/
tissues by the patient. Therapeutic compositions comprising cells/tissues
transformed with (I) is useful in identifying the role of genes in
healthy and diseased tissue, in tissue engineering and in cosmetic
surgery. Tissue engineering can be used to treat arthritis, joint
replacement, skin grafts for burn victims, and replacement coronary
arteries. Cosmetic tissue surgery includes rhinoplasty. The present
sequence represents the nucleotide sequence of the vector pIRES/BS
containing the BS gene
Sequence 4840 BP; 1154 A; 1227 C; 1236 G; 1223 T; 0 U; 0 Other;
Query Match 100.0%; Score 100; DB 4; Length 4840;
Best Local Similarity 100.0%; Pred. No. 5.6e-26;
Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTCATG 60
Db 1 GACGGATCGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAAATCTGCTCTCATG 60
Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
RESULT 15
ADB33528
ID ADB33528 standard; DNA; 5015 BP.
AC ADB33528;
XX
DT 04-DEC-2003 (first entry)
DE Expression vector nucleotide sequence SEQ ID NO:27.
XX
KW fusion protein; amyloid precursor protein; APP; transcription factor;
KW neurotrophic; neuroprotective; APP inhibitor;
KW amyloid precursor protein inhibitor; Alzheimer's disease; beta-secretase;
KW gamma-secretase; human; gene; ds.
XX
OS Synthetic.
OS Homo sapiens.
PN WO2003072041-A2.
XX
PD 04-SEP-2003.
XX
PF 23-FEB-2003; 2003WO-US0005458.

XX 27-FEB-2002; 2002US-0360274P.
 XX (MERI) MERCK & CO INC.
 XX Espeseth AS, Ferrer M, Flores OA, Hazuda DJ, Inglesse J;
 XX Miller MD, Register B, Shi X, Simon AJ, Zuck PD;
 XX WPI; 2003-689968/65.
 XX DNA encoding a fusion protein of amyloid precursor protein, useful in
 XX screening for anti-Alzheimer agents, comprises a fused transcription
 XX factor.
 XX Disclosure; Fig 32B-F; 193pp; English.
 XX The present invention describes a DNA molecule (I) that encodes a fusion
 XX protein (FP) comprising: (i) an amino acid sequence of amyloid precursor
 XX protein (APP), either the wild type, Swedish or NFEV versions; and (ii) a
 XX transcription factor (TF), fused in frame to the C-terminus of (i). Also
 XX described: (1) an expression vector containing (I); (2) a eukaryotic cell
 XX containing (I); and (3) methods for identifying a compound (A) that
 XX inhibits processing of APP, using the cells of (2). (I) has neurotropic and
 XX neuroprotective activities. (1) can be used to produce eukaryotic cells
 XX that express FP and are useful in screening for agents that inhibit
 XX processing of APP. The agents are potentially useful for the treatment or
 XX prevention of Alzheimer's disease. Cells that express FP can screen for
 XX inhibitors of: (a) beta- and gamma-secretases; and (b)
 XX cytoplasmic/extracellular APP signaling in a single assay. Cell-based
 XX assays may be free of interference from alpha-secretase activity and are
 XX homogeneous (no chromatography, immunoprecipitation or washing required)
 XX so well suited to high-throughput screening. The present sequence
 XX represents a plasmid nucleotide sequence from the present invention.
 XX SQ Sequence 5015 BP; 1167 A; 1297 C; 1279 G; 1272 T; 0 U; 0 Other;
 Query Match 100.0%; Score 100; DB 10; Length 5015;
 Best Local Similarity 100.0%; Pred. No. 5.6e-26;
 Matches 100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GACGATCGGAGATCCCGATCCCTATGCTGACTTCAGTACATCTGCTGATG 60
 Db 1 GACGATCGGAGATCCCGATCCCTATGCTGACTTCAGTACATCTGCTGATG 60
 Qy 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100
 Db 61 CCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTT 100

Search completed: July 14, 2005, 07:01:24
 Job time : 147.038 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6

Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 05:15:57 ; Search time 952.146 Seconds
(without alignments)
3997.736 Million cell updates/sec

Title: US-09-482-682-8_COPY_1_100

Perfect score: 100

Sequence: 1 gacggatcgagatctccc.....ctgtccctgtgtgtgtgt 100

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hic:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_ges1:*

9: gb_ges2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	60	60.0	602	8	B67169 CpG0047A Cp
2	55.6	55.6	694	8	B2052929 jnr13903.
3	55.6	55.6	696	8	B2050328 jnr42c12.
4	55.6	55.6	717	8	B2054067 jnr38b09.
5	53.6	53.6	348	2	AW409112 sal10h5 S
6	53.4	53.4	343	1	AL715724 AL715724
7	53.4	53.4	345	1	AL714571 AL714571
8	53.4	53.4	761	7	CK119397 212c09.pl
9	53.4	53.4	766	7	CK120360 207j04.pl
10	53.4	53.4	788	7	CK117844 209p08.pl
11	53.4	53.4	898	9	CK1141237 ISB1-118J
12	53.4	53.4	899	9	CK1140877 ISB1-118B
13	53.4	53.4	1009	9	CK123953 ISB1-84J1
14	53.2	53.2	814	8	AQ914559 nbe0049M
15	53	53.0	675	8	B2051815 jnr57d03.
16	53	53.0	679	8	B2052857 jnr13903.
17	53	53.0	700	8	B2050646 jnr66f08.
18	53	53.0	701	8	B2052015 jnr56b03.
19	53	53.0	708	8	B2054793 jnr33903.
20	53	53.0	709	8	B2053587 jnr98d01.
21	53	53.0	712	8	B2054005 jnr38b09.
22	52.8	52.8	451	8	AQ863966 nbe0022E
23	52.6	52.6	399	8	AQ075099 CIT-HSP-2
24	52.4	52.4	700	8	B2049113 jnr21d02.

25	52.4	52.4	708	8	B2050047
26	51.6	51.6	328	9	CC819886
27	51.6	51.6	351	9	CC818492
28	51.6	51.6	358	9	CC817661
29	51.6	51.6	364	9	CC817805
30	51.6	51.6	364	9	CC818511
31	51.6	51.6	364	9	CC818574
32	51.6	51.6	364	9	CC819049
33	51.6	51.6	369	9	CC817069
34	51.6	51.6	374	9	CC817074
35	51.6	51.6	374	9	CC820036
36	51.6	51.6	395	9	CC817652
37	51.6	51.6	403	9	CC817682
38	51.6	51.6	403	9	CC817837
39	51.6	51.6	414	9	CC819240
40	51.6	51.6	419	9	CC818384
41	51.6	51.6	420	9	CC817834
42	51.6	51.6	426	9	CC817720
43	51.6	51.6	437	9	CC819820
44	51.6	51.6	441	9	CC818421
45	51.6	51.6	443	9	CC817769

ALIGNMENTS

RESULT 1
B67169
LOCUS B67169 CpG0047A CpIOWAgDNA2 602 bp DNA linear GSS 12-MAY-2000
DEFINITION sequence.
ACCESSION B67169
VERSION B67169.1 GI:2642750
KEYWORDS GSS.
SOURCE Cryptosporidium parvum
ORGANISM Cryptosporidium parvum
REFERENCE Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
Cryptosporidiidae; Cryptosporidium.
1 (bases 1 to 602)
AUTHORS Strong, W.B. and Nelson, R.G.
TITLE Preliminary profile of the Cryptosporidium parvum genome: an
expressed sequence tag and genome survey sequence analysis
JOURNAL Mol. Biochem. Parasitol. 107 (1), 1-32 (2000)
MEDLINE 20183851
PUBMED 10717299
COMMENT Contact: Nelson, R. G.
Depts. of Medicine & Pharmaceutical Chemistry
San Francisco General Hospital-University of California, San
Francisco
Box 0811, San Francisco, CA 94143-0811, USA
Tel: 415 206 8846
Fax: 415 206 3353
Email: malaria@itsa.ucsf.edu
Submitted sequence has been edited to remove vector sequences 5' to
the insert, to correct miscalled bases and assign uncalled (N)
bases throughout the sequence, and to terminate when base-calling
became ambiguous.
Seq primer: T7
Class: shotgun
High quality sequence stop: 602.

FEATURES

source

1..602
Location/Qualifiers
/organism="Cryptosporidium parvum"
/mol_type="genomic DNA"
/strain="IOWA"
/db_xref="taxon:5807"
/lab_host="E. coli XL2 Blue MRF"
/clone_lib="CpIOWAgDNA2"
/note="Vector: PCR-Script Amp SK+; Site 1: SrfI; C. parvum
(IOWA isolate) genomic DNA was hydrodynamically sheared
to produce fragments having a tight size distribution
between 2-4 kb by Dr. Yvonne Thorstenson of the Stanford
DNA Sequencing and Technology Center

(http://sequence-www.stanford.edu/group/techdev/shear.htm)

The randomly sheared gDNA was chromatographed on Sephacryl S-400 to remove any small fragments and DNA eluting in the void volume was blunt-ended with T4 DNA Polymerase, treated with alkaline phosphatase to prevent the ligation of multiple fragments, ligated to Srf I-digested PCR-Script Amp (SK+) vector and transformed into E. coli strain XL10 Gold. Recombinant clones from the first plating of the library were selected for sequence analysis using T3 and T7 primers."

ORIGIN

Query Match 60.0%; Score 60; DB 8; Length 602;
Best Local Similarity 100.0%; Pred. No. 2.4e-10;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 41 CAGTACAAATCTGCTCTGATGCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 100
|||||
Db 1 CAGTACAAATCTGCTCTGATGCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTTGTT 60

RESULT 2

BZ052929/c
LOCUS jnr13g03.g1 B.oleracea001 Brassica oleracea genomic, genomic survey
DEFINITION sequence.

ACCESSION BZ052929
VERSION BZ052929.1 GI:23654922
KEYWORDS GSS.

SOURCE

ORGANISM Brassica oleracea

Brassica oleracea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

1 (bases 1 to 694)

Delehaunty, K., Fewell, G., Fulton, L., McCombie, W.R., Miner, T.,

Nash, W., Rabinowicz, P.D. and Wilson, R.K.

Whole genome shotgun reads from Brassica oleracea

Unpublished (2002)

Contact: Richard K. Wilson

Genome Sequencing Center

Washington University School of Medicine

Email: submissions@watson.wustl.edu

Plate: jnr13 row: 9 column: 03

Seq primer: -28RppOT reverse

Class: shotgun

High quality sequence start: 32

High quality sequence stop: 551.

FEATURES

source
Location/Qualifiers
1..694
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/db_xref="taxon:3712"

/clone_lib="B.oleracea001"
/note="Vector: pOTw13; Whole genome shotgun library from
flowering buds. DNA was purified from a crude nuclear
prep using Brassica oleracea T01000DH3 buds provided by
Thomas Osborn at the University of Wisconsin. Genomic
DNA was provided by Pablo Rabinowicz (CSHL) and the
shotgun library prepared at Washington University Genome
Sequencing Center."

ORIGIN

Query Match 55.6%; Score 55.6; DB 8; Length 694;
Best Local Similarity 77.9%; Pred. No. 9e-09;
Matches 67; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 3 CGGATCGGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATGCC 62
|||||
Db 324 CGGATCGATAGTCCCTCGGATAGTTATGTCGACTCTCAGTACAATCTGCTCTGATGCC 265

Qy

63 GCATAGTTAAACCGAGTATCTGCTCCC 88
|||||

Db 264 GCATAGTTAAGCCAGCCCGACACCC 239

RESULT 3

BZ050328

LOCUS

DEFINITION

sequence.

ACCESSION BZ050328

VERSION BZ050328.1 GI:23649718

KEYWORDS GSS.

SOURCE

ORGANISM Brassica oleracea

Brassica oleracea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

1 (bases 1 to 696)

Delehaunty, K., Fewell, G., Fulton, L., McCombie, W.R., Miner, T.,

Nash, W., Rabinowicz, P.D. and Wilson, R.K.

Whole genome shotgun reads from Brassica oleracea

Unpublished (2002)

Contact: Richard K. Wilson

Genome Sequencing Center

Washington University School of Medicine

Email: submissions@watson.wustl.edu

Plate: jnr42 row: C column: 12

Seq primer: -21UPpOT forward

Class: shotgun

High quality sequence start: 35

High quality sequence stop: 180.

FEATURES

source

Location/Qualifiers

1..696
/organism="Brassica oleracea"

/mol_type="genomic DNA"

/db_xref="taxon:3712"

/clone_lib="B.oleracea001"

/note="Vector: pOTw13; Whole genome shotgun library from

flowering buds. DNA was purified from a crude nuclear

prep using Brassica oleracea T01000DH3 buds provided by

Thomas Osborn at the University of Wisconsin. Genomic

DNA was provided by Pablo Rabinowicz (CSHL) and the

shotgun library prepared at Washington University Genome

Sequencing Center."

ORIGIN

Query Match 55.6%; Score 55.6; DB 8; Length 696;
Best Local Similarity 77.9%; Pred. No. 9e-09;
Matches 67; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 3 CGGATCGGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAATCTGCTCTGATGCC 62
|||||
Db 45 CGGATCGATAGTCCCTCGGATAGTTATGTCGACTCTCAGTACAATCTGCTCTGATGCC 104

Qy

63 GCATAGTTAAGCCAGTATCTGCTCCC 88
|||||

Db 105 GCATAGTTAAGCCAGCCCGACACCC 130

RESULT 4

BZ054067/c

LOCUS

DEFINITION

sequence.

ACCESSION BZ054067

VERSION BZ054067.1 GI:23657216

KEYWORDS GSS.

SOURCE

ORGANISM Brassica oleracea

Brassica oleracea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Brassica.

1 (bases 1 to 717)

Delehaunty, K., Fewell, G., Fulton, L., McCombie, W.R., Miner, T.,

TITLE
JOURNAL
COMMENT

Nash, W., Rabinowicz, P.D. and Wilson, R.K.
Whole genome shotgun reads from Brassica oleracea
Unpublished (2002)
Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Plate: jnr38 row: b column: 09
Seq primer: -28RppOT reverse
Class: shotgun
High quality sequence start: 87
High quality sequence stop: 543.

FEATURES

source
Location/Qualifiers
1. .717
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/db_xref="taxon:3712"
/clone_lib="B.oleracea001"
/note="Vector: pOTW13; Whole genome shotgun library from
flowering buds. DNA was purified from a crude nuclear
prep using Brassica oleracea T01000DH3 buds provided by
Thomas Osborn at the University of Wisconsin. Genomic
DNA was provided by Pablo Rabinowicz (CSHL) and the
shotgun library prepared at Washington University Genome
Sequencing Center."

ORIGIN

Query Match 55.6%; Score 55.6; DB 8; Length 717;
Best Local Similarity 77.9%; Pred. No. 9.1e-09;
Matches 67; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
Qy 3 CGGATCGGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAACTGCTGTGATGCC 62
Db 248 CGGATCGATAGTCCCTGGATAGTATGTCGACTCTCAGTACAACTGCTGTGATGCC 189
Qy 63 GCATAGTTAAGCCAGTATCTGCTCC 88
Db 188 GCATAGTTAAGCCAGCCCGACACC 163

ORIGIN

Query Match 55.6%; Score 55.6; DB 8; Length 717;
Best Local Similarity 77.9%; Pred. No. 9.1e-09;
Matches 67; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
Qy 3 CGGATCGGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAACTGCTGTGATGCC 62
Db 248 CGGATCGATAGTCCCTGGATAGTATGTCGACTCTCAGTACAACTGCTGTGATGCC 189
Qy 63 GCATAGTTAAGCCAGTATCTGCTCC 88
Db 188 GCATAGTTAAGCCAGCCCGACACC 163

ORIGIN

Query Match 55.6%; Score 55.6; DB 8; Length 717;
Best Local Similarity 77.9%; Pred. No. 9.1e-09;
Matches 67; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
Qy 3 CGGATCGGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAACTGCTGTGATGCC 62
Db 248 CGGATCGATAGTCCCTGGATAGTATGTCGACTCTCAGTACAACTGCTGTGATGCC 189
Qy 63 GCATAGTTAAGCCAGTATCTGCTCC 88
Db 188 GCATAGTTAAGCCAGCCCGACACC 163

ORIGIN

Query Match 55.6%; Score 55.6; DB 8; Length 717;
Best Local Similarity 77.9%; Pred. No. 9.1e-09;
Matches 67; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
Qy 3 CGGATCGGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAACTGCTGTGATGCC 62
Db 248 CGGATCGATAGTCCCTGGATAGTATGTCGACTCTCAGTACAACTGCTGTGATGCC 189
Qy 63 GCATAGTTAAGCCAGTATCTGCTCC 88
Db 188 GCATAGTTAAGCCAGCCCGACACC 163

ORIGIN

Query Match 55.6%; Score 55.6; DB 8; Length 717;
Best Local Similarity 77.9%; Pred. No. 9.1e-09;
Matches 67; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
Qy 3 CGGATCGGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAACTGCTGTGATGCC 62
Db 248 CGGATCGATAGTCCCTGGATAGTATGTCGACTCTCAGTACAACTGCTGTGATGCC 189
Qy 63 GCATAGTTAAGCCAGTATCTGCTCC 88
Db 188 GCATAGTTAAGCCAGCCCGACACC 163

ORIGIN

Query Match 55.6%; Score 55.6; DB 8; Length 717;
Best Local Similarity 77.9%; Pred. No. 9.1e-09;
Matches 67; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
Qy 3 CGGATCGGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAACTGCTGTGATGCC 62
Db 248 CGGATCGATAGTCCCTGGATAGTATGTCGACTCTCAGTACAACTGCTGTGATGCC 189
Qy 63 GCATAGTTAAGCCAGTATCTGCTCC 88
Db 188 GCATAGTTAAGCCAGCCCGACACC 163

ORIGIN

Query Match 55.6%; Score 55.6; DB 8; Length 717;
Best Local Similarity 77.9%; Pred. No. 9.1e-09;
Matches 67; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
Qy 3 CGGATCGGGAGATCTCCGATCCCTATGTCGACTCTCAGTACAACTGCTGTGATGCC 62
Db 248 CGGATCGATAGTCCCTGGATAGTATGTCGACTCTCAGTACAACTGCTGTGATGCC 189
Qy 63 GCATAGTTAAGCCAGTATCTGCTCC 88
Db 188 GCATAGTTAAGCCAGCCCGACACC 163

Query Match 53.6%; Score 53.6; DB 2; Length 348;
Best Local Similarity 80.5%; Pred. No. 4.1e-08;
Matches 62; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 11 GAGATCTCCGATCCCTATGTCGACTCTCAGTACAACTGCTGTGATGCCGATAGTT 70
Db 65 GCGGTATACACCGCATATGTCGACTCTCAGTACAACTGCTGTGATGCCGATAGTT 124

Qy 71 AAGCCAGTATCTGCTCC 87
Db 125 AAGCCAGTATACACTCC 141

FEATURES

source
Location/Qualifiers
1. .343
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="BN0AA018ZF12"
/tissue_type="inner ear"
/dev_stage="embryonic"
/clone_lib="Danio rerio embryonic inner ear subtracted
cdna"
/note="subtracted cdna library"

AL715724 343 bp mRNA linear EST 18-APR-2002
DEFINITION AL715724 Danio rerio embryonic inner ear subtracted CDNA Danio
rerio CDNA clone BN0AA018ZF12 5', mRNA sequence.
ACCESSION AL715724.1 GI:20180327
VERSION AL715724.1
KEYWORDS EST.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.

REFERENCE 1 (bases 1 to 343)
AUTHORS Coimbra, R., Weil, D., Brottier, P., Blanchard, S., Levi, M.,
Hardell, J.P., Weissenbach, J. and Petit, C.
TITLE A subtracted cDNA library from the zebrafish (Danio rerio)
embryonic inner ear
JOURNAL Unpublished (2002)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
Location/Qualifiers
1. .343
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="BN0AA018ZF12"
/tissue_type="inner ear"
/dev_stage="embryonic"
/clone_lib="Danio rerio embryonic inner ear subtracted
cdna"
/note="subtracted cdna library"

AL715724 343 bp mRNA linear EST 18-APR-2002
DEFINITION AL715724 Danio rerio embryonic inner ear subtracted CDNA Danio
rerio CDNA clone BN0AA018ZF12 5', mRNA sequence.
ACCESSION AL715724.1 GI:20180327
VERSION AL715724.1
KEYWORDS EST.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.

REFERENCE 1 (bases 1 to 343)
AUTHORS Coimbra, R., Weil, D., Brottier, P., Blanchard, S., Levi, M.,
Hardell, J.P., Weissenbach, J. and Petit, C.
TITLE A subtracted cDNA library from the zebrafish (Danio rerio)
embryonic inner ear
JOURNAL Unpublished (2002)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
Location/Qualifiers
1. .343
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="BN0AA018ZF12"
/tissue_type="inner ear"
/dev_stage="embryonic"
/clone_lib="Danio rerio embryonic inner ear subtracted
cdna"
/note="subtracted cdna library"

AL715724 343 bp mRNA linear EST 18-APR-2002
DEFINITION AL715724 Danio rerio embryonic inner ear subtracted CDNA Danio
rerio CDNA clone BN0AA018ZF12 5', mRNA sequence.
ACCESSION AL715724.1 GI:20180327
VERSION AL715724.1
KEYWORDS EST.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.

REFERENCE 1 (bases 1 to 343)
AUTHORS Coimbra, R., Weil, D., Brottier, P., Blanchard, S., Levi, M.,
Hardell, J.P., Weissenbach, J. and Petit, C.
TITLE A subtracted cDNA library from the zebrafish (Danio rerio)
embryonic inner ear
JOURNAL Unpublished (2002)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
Location/Qualifiers
1. .343
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="BN0AA018ZF12"
/tissue_type="inner ear"
/dev_stage="embryonic"
/clone_lib="Danio rerio embryonic inner ear subtracted
cdna"
/note="subtracted cdna library"

AL715724 343 bp mRNA linear EST 18-APR-2002
DEFINITION AL715724 Danio rerio embryonic inner ear subtracted CDNA Danio
rerio CDNA clone BN0AA018ZF12 5', mRNA sequence.
ACCESSION AL715724.1 GI:20180327
VERSION AL715724.1
KEYWORDS EST.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.

REFERENCE 1 (bases 1 to 343)
AUTHORS Coimbra, R., Weil, D., Brottier, P., Blanchard, S., Levi, M.,
Hardell, J.P., Weissenbach, J. and Petit, C.
TITLE A subtracted cDNA library from the zebrafish (Danio rerio)
embryonic inner ear
JOURNAL Unpublished (2002)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
Location/Qualifiers
1. .343
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="BN0AA018ZF12"
/tissue_type="inner ear"
/dev_stage="embryonic"
/clone_lib="Danio rerio embryonic inner ear subtracted
cdna"
/note="subtracted cdna library"

```

Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 345)
AUTHORS      Coimbra,R., Wei,D., Brottier,P., Blanchard,S., Levi,M.,
              Hardelin,J.P., Weissenbach,J. and Petit,C.
TITLE        A subtracted cDNA library from the zebrafish (Danio rerio)
              embryonic inner ear
JOURNAL      Unpublished (2002)
COMMENT      Contact: Genoscope
              Genoscope - Centre National de Sequencage
              2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
              Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES     source
              1..345
               /organism="Danio rerio"
               /mol_type="mRNA"
               /db_xref="taxon:7955"
               /clone="BN0AA0072C02"
               /tissue_type="inner ear"
               /dev_stage="embryonic"
               /clone_lib="Danio rerio embryonic inner ear subtracted
               cDNA"
               /note="subtracted cDNA library"

ORIGIN
Query Match      53.4%; Score 53.4; DB 1; Length 345;
Best Local Similarity 84.5%; Pred. No. 4.8e-08;
Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 17 TCCCGATCCCTATGTCGACTCTCAGTACAACTCTGCTGTGATGCCGATAGTTAAGCCA 76
DB 280 TTACACCGCATATGGTGCACCTCTCAGTACAACTCTGCTGTGATGCCGATAGTTAAGCCA 221

QY 77 GTATCTGCTCC 87
DB 220 GTATACACTCC 210

RESULT 8
CK119397/c
LOCUS      CK119397
DEFINITION 212009.p1 AtM1 Arabidopsis thaliana cDNA clone MPMGp2011009212
5-PRIME, mRNA sequence.
ACCESSION  CK119397
VERSION     CK119397.1 GI:47829713
KEYWORDS    EST.
SOURCE      Arabidopsis thaliana (thale cress)
ORGANISM    Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 761)
Feilner,T., Immink,R.G.H., Cahill,D.J. and Kersten,B.
Generation of a cDNA expression library from Arabidopsis
inflorescence meristem
Unpublished (2003)
Contact: Birgit Kersten
Plant Protein Chip Group, Department Lehrach
Max-Planck-Institute for Molecular Genetics
Innestr. 73 , D-14195 Berlin, Germany
Tel: +49(0)30/84131648
Fax: +49(0)30/84131128
Email: Kersten@molgen.mpg.de
Insert Length: 761 Std Error: 0.00
Plate: 212 row: 0 column: 9
Seq primer: PQB65.
              Location/Qualifiers
              1..761
               /organism="Arabidopsis thaliana"
               /mol_type="mRNA"
               /ecotype="Columbia"
               /db_xref="GABI:954234"
               /db_xref="taxon:3702"

REFERENCE
AUTHORS      Feilner,T., Immink,R.G.H., Cahill,D.J. and Kersten,B.
TITLE        Generation of a cDNA expression library from Arabidopsis
              inflorescence meristem
JOURNAL      Unpublished (2003)
COMMENT      Contact: Birgit Kersten
              Plant Protein Chip Group, Department Lehrach
              Max-Planck-Institute for Molecular Genetics
              Innestr. 73 , D-14195 Berlin, Germany
              Tel: +49(0)30/84131648
              Fax: +49(0)30/84131128
              Email: Kersten@molgen.mpg.de
              Insert Length: 766 Std Error: 0.00
              Plate: 207 row: J column: 4
              Seq primer: PQB65.
              Location/Qualifiers
              1..766
               /organism="Arabidopsis thaliana"
               /mol_type="mRNA"
               /ecotype="Columbia"
               /db_xref="GABI:953059"
               /db_xref="taxon:3702"
               /clone="MPMGp2011J04207"
               /tissue_type="inflorescence meristem"
               /dev_stage="about one week after bolting"
               /lab_host="E. coli SCS-1/pSE111"
               /clone_lib="AtM1"
               /note="Vector: PQE-3ONAST-attB (AY386205); Site_1: SalI;

```

```

/clone="MPMGp2011009212"
/tissue_type="inflorescence meristem"
/dev_stage="about one week after bolting"
/lab_host="E. coli SCS-1/pSE111"
/clone_lib="AtM1"
/note="Vector: PQE-3ONAST-attB (AY386205); Site_1: SalI;
Site_2: NotI; About 1 week after bolting, cDNA synthesis
using SuperscriptTM-system (Invitrogen) with an
oligo(dT)-primer containing NotI restriction site and a
SalI adapter. The main library (plate numbers begin with
1) of 38,000 clones was rearrayed into the sublibrary
(plate numbers begin with 201) containing 5,000 putative
expression clones. Average insert size is 1 kb. Note: The
rearrayed sublibrary (plate numbers begin with 201) was
sequenced. Library generation and sequencing was granted
in context of GABI-LAPP; data are also accessible at
https://gabi.rzpd.de"

ORIGIN
Query Match      53.4%; Score 53.4; DB 7; Length 761;
Best Local Similarity 84.5%; Pred. No. 5.6e-08;
Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 17 TCCCGATCCCTATGTCGACTCTCAGTACAACTCTGCTGTGATGCCGATAGTTAAGCCA 76
DB 674 TTACACCGCATATGGTGCACCTCTCAGTACAACTCTGCTGTGATGCCGATAGTTAAGCCA 615

QY 77 GTATCTGCTCC 87
DB 614 GTATACACTCC 604

RESULT 9
CK120360/c
LOCUS      CK120360
DEFINITION 207J04.p1 AtM1 Arabidopsis thaliana cDNA clone MPMGp2011J04207
5-PRIME, mRNA sequence.
ACCESSION  CK120360
VERSION     CK120360.1 GI:47830676
KEYWORDS    EST.
SOURCE      Arabidopsis thaliana (thale cress)
ORGANISM    Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 766)
Feilner,T., Immink,R.G.H., Cahill,D.J. and Kersten,B.
Generation of a cDNA expression library from Arabidopsis
inflorescence meristem
Unpublished (2003)
Contact: Birgit Kersten
Plant Protein Chip Group, Department Lehrach
Max-Planck-Institute for Molecular Genetics
Innestr. 73 , D-14195 Berlin, Germany
Tel: +49(0)30/84131648
Fax: +49(0)30/84131128
Email: Kersten@molgen.mpg.de
Insert Length: 766 Std Error: 0.00
Plate: 207 row: J column: 4
Seq primer: PQB65.
              Location/Qualifiers
              1..766
               /organism="Arabidopsis thaliana"
               /mol_type="mRNA"
               /ecotype="Columbia"
               /db_xref="GABI:953059"
               /db_xref="taxon:3702"
               /clone="MPMGp2011J04207"
               /tissue_type="inflorescence meristem"
               /dev_stage="about one week after bolting"
               /lab_host="E. coli SCS-1/pSE111"
               /clone_lib="AtM1"
               /note="Vector: PQE-3ONAST-attB (AY386205); Site_1: SalI;

```

expression clones. Average insert size is 1 kb. Note: The rearranged sublibrary (plate numbers begin with 201) was sequenced. Library generation and sequencing was granted in context of GABI-LAPP; data are also accessible at <https://gabi.rzpd.de>

Db 514 TTACACCGCATATGGTGCACTCTCAGTACAATCTGTCTGATGCCGATAGTTAAGCA 455

QY 77 GTATCTGCTCC 87

Db 454 GTATACACTCC 444

CL141237
 ACCESSION
 VERSION
 CL141237.1
 GI:40634872
 CDS
 FEATURES

Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae; Xenohipodinae; Xenopus; Silurana.
1 (bases 1 to 898)

REFERENCE
AUTHORS
Kremitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.
I (Dues 1 to 050)

THE JOURNAL OF THE PHYSICAL SCIENCES
PUBLISHED MONTHLY (2003)

Genome Sequencing Center
Washington University School of Medicine

```
INVERT Length: 75000 STD Error: 0.00  
SOS max: 77 MAXVAL: 99999999
```

High quality sequence start: 4
High quality sequence stop: 743

source
1. .898
/count="yes";

```
/db_xref="taxon:8364"
```

/note="Vector: pBelovBAC11; ISB-1 *Xenopus tropicalis* BAC

Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Qy 77 GTATCTGCTCC 87

CL140877/c

```

LOCUS      CL140877                      899 bp    DNA        linear    GSS 05-JAN-2004
DEFINITION ISB1-118B12.T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-118B12,
genomic survey sequence.
ACCESSION  CL140877
VERSION    CL140877.1 GI:40634512
KEYWORDS   GSS.
SOURCE     Xenopus tropicalis (western clawed frog)
ORGANISM   Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE  1 (bases 1 to 899)
AUTHORS    Kremitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
            Mardis, E. and Wilson, R.
TITLE      A physical map of the xenopus tropicalis genome
JOURNAL    Unpublished (2003)
COMMENT    Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 75000 Std Error: 0.00
            Seq primer: T7 TAATACGACTCACTATAGG
            Class: BAC ends
            High quality sequence start: 4
            High quality sequence stop: 681.
FEATURES   source
            Location/Qualifiers
                1..899
                /organism="Xenopus tropicalis"
                /mol_type="genomic DNA"
                /db_xref="taxon:8364"
                /clone="ISB1-118B12"
                /clone_lib="ISB1"
                /note="vector: phelOBAC11; ISB-1 Xenopus tropicalis BAC
                Library Segment 1"
ORIGIN
Query Match      53.4%; Score 53.4; DB 9; Length 899;
Best Local Similarity 84.5%; Pred. No. 5.8e-08;
Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 17 TCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTGATGCCCGCATAGTTAAGCCA 76
Db 195 TTACACGGCATATGTCGACTCTCAGTACAATCTGCTGATGCCCGCATAGTTAAGCCA 136
QY 77 GTATCTGCTCC 87
Db 135 GTATACACTCC 125

RESULT 13
LOCUS      CL123953/c
DEFINITION ISB1-84J15.T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-84J15,
genomic survey sequence.
ACCESSION  CL123953
VERSION    CL123953.1 GI:40617598
KEYWORDS   GSS.
SOURCE     Xenopus tropicalis (western clawed frog)
ORGANISM   Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE  1 (bases 1 to 1009)
AUTHORS    Kremitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
            Mardis, E. and Wilson, R.
TITLE      A physical map of the xenopus tropicalis genome
JOURNAL    Unpublished (2003)
COMMENT    Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 75000 Std Error: 0.00
            Seq primer: T7 TAATACGACTCACTATAGG

```

```

Class: BAC ends
High quality sequence start: 167
High quality sequence stop: 324.
FEATURES   source
            Location/Qualifiers
                1..1009
                /organism="Xenopus tropicalis"
                /mol_type="genomic DNA"
                /db_xref="taxon:8364"
                /clone="ISB1-84J15"
                /clone_lib="ISB1"
                /note="vector: pBeloBAC11; ISB-1 Xenopus tropicalis BAC
                Library Segment 1"
ORIGIN
Query Match      53.4%; Score 53.4; DB 9; Length 1009;
Best Local Similarity 84.5%; Pred. No. 5.9e-08;
Matches 60; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 17 TCCCGATCCCTATGTCGACTCTCAGTACAATCTGCTGATGCCCGCATAGTTAAGCCA 76
Db 252 TTACACGGCATATGTCGACTCTCAGTACAATCTGCTGATGCCCGCATAGTTAAGCCA 193
QY 77 GTATCTGCTCC 87
Db 192 GTATACACTCC 182

RESULT 14
LOCUS      AQ914559
DEFINITION nbe0049M21r CUGI Rice BAC Library (EcoRI) Oryza sativa (japonica
cultivar-group) genomic clone nbe0049M21r, genomic survey
sequence.
ACCESSION  AQ914559
VERSION    AQ914559.1 GI:6511075
KEYWORDS   GSS.
SOURCE     Oryza sativa (japonica cultivar-group)
            Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzaceae; Oryza.
            1 (bases 1 to 814)
            Wing, R.A. and Dean, R.A.
            A BAC End Sequencing Framework to Sequence the Rice Genome
            Unpublished (1998)
            Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Seq primer: GGAACAGCTATGACCATG
            Class: BAC ends
            High quality sequence start: 21
            High quality sequence stop: 361.
FEATURES   source
            Location/Qualifiers
                1..814
                /organism="Oryza sativa (japonica cultivar-group)"
                /mol_type="genomic DNA"
                /cultivar="Nipponbare"
                /db_xref="taxon:39947"
                /clone="nbe0049M21r"
                /tissue_type="Leaf"
                /lab_host="E. coli DH10B"
                /clone_lib="CUGI Rice BAC Library (EcoRI)"
                /note="vector: pBACIndigo; Site_1: EcoRI; Site_2: EcoRI;
                Rice is the most important food crop in the world. Half of
                the world population, especially those inhabiting highly
                populated areas of the humid tropics and subtropics, rely
                on rice as their primary source of carbohydrate.
                Monocotyledonous rice is a diploid plant (2n=24) with a
                haploid genome equivalent of 431 Mbp (Arumuganathan and

```

ORIGIN

RESULT 15

TITLE

FEATURES

ORIGIN

Query Match	53.0%	Score 53;	DB 8;	Length 675;
Best Local Similarity	75.6%	Pred. No. 7.6e-08;		

Search completed: July 14, 2005, 23:22:24
Job time : 957.146 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:39:07 ; Search time 756.618 Seconds
(without alignments)
6468.225 Million cell updates/sec

Title: US-09-482-682-8_COPY_3674_3774
Perfect score: 101
Sequence: 1 CGGTCTATTCTTTGATTATTAAGGGATTTCGGGATTTCGCCCTATTGGTTAAAAAATG 60

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101	100.0	3986	12	PCDNA3ZEO
2	101	100.0	4597	6	AX060344 Sequence
3	101	100.0	5070	6	AX234391 Sequence
4	101	100.0	5082	6	A91754 Sequence 10
5	101	100.0	5082	6	BD085110
6	101	100.0	5432	6	BD234590
7	101	100.0	5432	6	AX026821 Screening
8	101	100.0	5446	6	BD195386
9	101	100.0	5446	6	AX319694
10	101	100.0	5639	12	AX437643
11	101	100.0	5651	6	AX211282
12	101	100.0	5651	6	AX349366 Sequence
13	101	100.0	5731	6	AX202478
14	101	100.0	5995	6	AX685746
15	101	100.0	6084	12	GGA575208
16	101	100.0	6109	12	TRU90717
17	101	100.0	6148	6	BD181637
18	101	100.0	6148	6	AX342685
19	101	100.0	6149	6	BD181638

20	101	100.0	6149	6	AX342686	AX342686 Sequence
21	101	100.0	6180	6	AX207724	AX207724 Sequence
22	101	100.0	6186	6	AX211281	AX211281 Sequence
23	101	100.0	6186	6	AX349365	AX349365 Sequence
24	101	100.0	6195	6	BD168975	BD168975 Method of
25	101	100.0	6213	6	AX211283	AX211283 Sequence
26	101	100.0	6213	6	AX349369	AX349369 Sequence
27	101	100.0	6232	6	AR409004	AR409004 Sequence
28	101	100.0	6238	6	BD168966	BD168966 Method of
29	101	100.0	6253	6	AR031374	AR031374 Sequence
30	101	100.0	6253	6	BD009742	BD009742 Compositi
31	101	100.0	6277	12	AY437644	AY437644 Expressio
32	101	100.0	6331	12	EVPCMVPA1	X96612 Expression
33	101	100.0	6333	12	EVPCMVPA3	X96611 Expression
34	101	100.0	6335	12	EVPCMVPA2	X96610 Expression
35	101	100.0	6338	6	BD134374	BD134374 Peptide 1
36	101	100.0	6338	6	AR428934	AR428934 Sequence
37	101	100.0	6340	6	AX207733	AX207733 Sequence
38	101	100.0	6365	6	AX513181	AX513181 Sequence
39	101	100.0	6394	12	AF416990	AF416990 Synthetic
40	101	100.0	6404	6	BD267665	BD267665 Delivery
41	101	100.0	6411	6	AX207725	AX207725 Sequence
42	101	100.0	6411	6	AX207729	AX207729 Sequence
43	101	100.0	6420	6	BD267666	BD267666 Delivery
44	101	100.0	6436	6	AX207740	AX207740 Sequence
45	101	100.0	6439	6	AR240214	AR240214 Sequence

ALIGNMENTS

RESULT 1
PCDNA3ZEO
LOCUS PCDNA3ZEO 3986 bp DNA linear SYN 16-AUG-1995
DEFINITION Cloning vector pCDNA3ZEO DNA.
ACCESSION X90639
VERSION X90639.1 GI:949972
KEYWORDS cloning vector; expression vector; multiple cloning site; Plasmid.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Peters, H., Hundhausen, T., Kroenke, M. and Marget, M.
TITLE A new small sized high-level eukaryotic expression vector
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3986)
AUTHORS Peters, H.
TITLE Direct Submission
JOURNAL Submitted (07-AUG-1995): H. Peters, Inst. f. Immunologie,
Michaelisstr. 5, D- 24105 Kiel, FRG
COMMENT Related sequences: M21295 and K03104.
FEATURES
source
Location/Qualifiers
1..3986
/organism="synthetic construct"
/mol_type="other DNA"
/db_xref="taxon:32630"
/plasmid="pcDNA3ZEO"
1..2125
misc_feature /note="cloning vector (pcDNA3). (Invitrogen)"
misc_feature 889..994 /note="multiple cloning site (MCS)"
misc_feature 2126..2796 /note="cloning vector (PZeoSV) (Invitrogen)"
misc_feature 2797..3986 /note="cloning vector (pcDNA3)"

ORIGIN

Query Match 100.0%; Score 101; DB 12; Length 3986;
Best Local Similarity 100.0%; Pred. No. 5.1e-15;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGGTCTATTCTTTGATTATTAAGGGATTTCGGGATTTCGCCCTATTGGTTAAAAAATG 60
|||||

Db 1651 CGGTCATCTTTGATTATATAAGGATTTGGGGATTTTCGGCCTATTGGTTAAAAAATG 1710

Qy 61 AGCTGATTAAACAAAAATTTAACGCGAATTAATCTGTGGA 101
|||||
Db 1711 AGCTGATTAAACAAAAATTTAACGCGAATTAATCTGTGGA 1751

RESULT 2
AX060344
LOCUS AX060344 4597 bp DNA linear PAT 22-JAN-2001
DEFINITION Sequence 3 from Patent WO0078358.
ACCESSION AX060344
VERSION AX060344.1 GI:12405832
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Chen, W.
TITLE Hyaluronic acid microspheres for sustained gene transfer
JOURNAL Patent: WO 0078358-A 3 28-DEC-2000;
The Collaborative Group, Ltd. (US)
FEATURES
source
1. 4597
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="pCDNA3.1/GS vector by Invitrogen Corporation"

ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 4597;
Best Local Similarity 100.0%; Pred. No. 4.9e-15;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGTCATCTTTGATTATATAAGGATTTGGGGATTTTCGGCCTATTGGTTAAAAAATG 60
|||||

Db 2227 CGGTCATCTTTGATTATATAAGGATTTGGGGATTTTCGGCCTATTGGTTAAAAAATG 2286
|||||

Qy 61 AGCTGATTAAACAAAAATTTAACGCGAATTAATCTGTGGA 101
|||||

Db 2287 AGCTGATTAAACAAAAATTTAACGCGAATTAATCTGTGGA 2327
|||||

RESULT 3
AX234391
LOCUS AX234391 5070 bp DNA linear PAT 11-SEP-2001
DEFINITION Sequence 41 from Patent WO0162942.
ACCESSION AX234391
VERSION AX234391.1 GI:15593392
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Altalo, K.M. and Jeltsch, M.M.
TITLE Materials and methods involving hybrid vascular endothelial growth factor dnas and proteins and screening methods for modulators
JOURNAL Patent: WO 0162942-A 41 30-AUG-2001;
LUDWIG INSTITUTE FOR CANCER RESEARCH (US); Licentia OY (FI)
FEATURES
source
1. 5070
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="pSectagI Vector"

ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 5070;
Best Local Similarity 100.0%; Pred. No. 4.8e-15;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGTCATCTTTGATTATATAAGGATTTGGGGATTTTCGGCCTATTGGTTAAAAAATG 60
|||||

Db 1693 CGGTCATCTTTGATTATATAAGGATTTGGGGATTTTCGGCCTATTGGTTAAAAAATG 1752

Qy 61 AGCTGATTAAACAAAAATTTAACGCGAATTAATCTGTGGA 101
|||||
Db 1753 AGCTGATTAAACAAAAATTTAACGCGAATTAATCTGTGGA 1793
|||||

RESULT 4
A91754
LOCUS A91754 5082 bp DNA circular PAT 22-JAN-2000
DEFINITION Sequence 10 from Patent WO9824810.
ACCESSION A91754
VERSION A91754.1 GI:6740671
KEYWORDS unidentified
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 5082)
AUTHORS Bogaert, T.A. and Deraeymaeker, M.
TITLE VERTEBRATE HOMOLOGUES OF UNC-53 PROTEIN OF C. ELEGANS
JOURNAL Patent: WO 9824810-A 10 11-JUN-1998;
BOGAERT THIERRY ANDRE OLIVIER (BE); DERAEMYAEKER MARC (BE)
FEATURES
Location/Qualifiers
source
1. 5082
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 5082;
Best Local Similarity 100.0%; Pred. No. 4.8e-15;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGTCATCTTTGATTATATAAGGATTTGGGGATTTTCGGCCTATTGGTTAAAAAATG 60
|||||

Db 3321 CGGTCATCTTTGATTATATAAGGATTTGGGGATTTTCGGCCTATTGGTTAAAAAATG 3380
|||||

Qy 61 AGCTGATTAAACAAAAATTTAACGCGAATTAATCTGTGGA 101
|||||

Db 3381 AGCTGATTAAACAAAAATTTAACGCGAATTAATCTGTGGA 3421
|||||

RESULT 5
BD085110
LOCUS BD085110 5082 bp DNA linear PAT 27-AUG-2002
DEFINITION Vertebrate homologues of UNC-53 protein of C elegans.
ACCESSION BD085110
VERSION BD085110.1 GI:22630720
KEYWORDS JP 2001522222-A/8.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 5082)
AUTHORS Platteuw, C.J., Arjol, C.M.B., Deraeymaeker, M., Verhasselt, P., Pujol, N.J.R., Luc, Maertens, J.S., Luyten, W., Geerts, H., Vandekerckhove, J.S., Geysen, J. and Bogaert, T.A.O.E.
TITLE Vertebrate homologues of UNC-53 protein of C elegans
JOURNAL Patent: JP 2001522222-A 8 13-NOV-2001;
JANSSEN PHARMACEUTICA NV
FEATURES
Location/Qualifiers
source
OS Unidentified
PN JP 2001522222-A/8
PD 13-NOV-2001
PF 03-DEC-1997 JP 1998525231
PR 04-DEC-1996 GB 9625283.8
PI CHRIST JULES PLATTEUW, CARLOS MANUEL BUESA ARJOL, MARC PI DERAEMYAEKER,
PI PETER VERHASSELT, NATHALIE JEANNE RAYMONDE PUJOL, LUC PI JACQUES SIMON MAERTENS,
PI WALTER LUYTEN, HUGO GEERTS, JOEL STEFAAN VANDEKERCKHOVE, JOHAN PI 'GEYSEN',
PI THIERRY ANDRE OLIVIER EDDY BOGAERT
PC C12N15/12, C12N5/10, C12N15/85, C07K14/435, C07K16/18, A61K38/17, A61K49/00,

```
PC C12Q1/02.G01N33/53
CC Strandedness: Double;
CC Topology: Circular;
CC Vertebrate homologues of UNC-53 protein of C elegans FH Key
FT source 1..5082 /organism='Unidentified'.
FT Location/Qualifiers
FEATURES
    source
        1..5082
        /organism='unidentified'
        /mol_type='genomic DNA'
        /db_xref='taxon:32644'
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 5082;
Best Local Similarity 100.0%; Pred. No. 4.8e-15;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGGTCATCTTTTGATTATTAAGGATTTGGGGATTTGGCCCTATTGGTTAAAAAATG 60
Db 3321 CGGTCATCTTTTGATTATTAAGGATTTGGGGATTTGGCCCTATTGGTTAAAAAATG 3380
QY 61 AGCTGATTTACAAAAATTTAACGCGAATTAATCTGTGGA 101
Db 3381 AGCTGATTTACAAAAATTTAACGCGAATTAATCTGTGGA 3421
RESULT 6
LOCUS BD234590 5432 bp DNA linear PAT 17-JUL-2003
DEFINITION Screening assay of Abeta-peptide.
ACCESSION BD234590
VERSION BD234590.1 GI:33044360
KEYWORDS JP 2002531141-A/2.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE
    1 (bases 1 to 5432)
    Peraus, G.
    Screening assay of Abeta-peptide
    Patent: JP 2002531141-A 2 24-SEP-2002;
    AVENTIS PHARMA DEUTSCHLAND GMBH
    OS Artificial Sequence
    PN JP 2002531141-A/2
    PD 24-SEP-2002
    PF 27-NOV-1999 JP 2000586944
    PR 07-DEC-1998 DE 198 56 261.6
    PI GISELA PERAUS
    PC C12N15/09, A01K67/033, A61K45/00, A61P25/28, C12N1/15, C12N1/19, PC
    C12N1/21,
    PC C12N5/10, C12Q1/37, C12Q1/68, C12N15/00, C12N5/00 CC Description
    of Artificial Sequence: Mutagen
    FT Key Location/Qualifiers
    FT source 1..5432
    FT Location/Qualifiers
    FT /organism='Artificial Sequence'.
FEATURES
    source
        1..5432
        /organism='synthetic construct'
        /mol_type='genomic DNA'
        /db_xref='taxon:32630'
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 5432;
Best Local Similarity 100.0%; Pred. No. 4.7e-15;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGGTCATCTTTTGATTATTAAGGATTTGGGGATTTGGCCCTATTGGTTAAAAAATG 60
Db 1637 CGGTCATCTTTTGATTATTAAGGATTTGGGGATTTGGCCCTATTGGTTAAAAAATG 1696
QY 61 AGCTGATTTACAAAAATTTAACGCGAATTAATCTGTGGA 101
Db 1697 AGCTGATTTACAAAAATTTAACGCGAATTAATCTGTGGA 1737
RESULT 7
LOCUS AX026821 5432 bp DNA linear PAT 16-SEP-2000
DEFINITION Sequence 9 from Patent DE19856261.
ACCESSION AX026821
VERSION AX026821.1 GI:10187947
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE
    1
    Peraus, G.
    Patent: DE 19856261-C 9 30-MAR-2000;
    HOECHST MARION ROUSSEL DE GMBH (DE)
    Location/Qualifiers
    1..5432
    /organism='synthetic construct'
    /mol_type='unassigned DNA'
    /db_xref='taxon:32630'
    /note='Mutagen'
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 5432;
Best Local Similarity 100.0%; Pred. No. 4.7e-15;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGGTCATCTTTTGATTATTAAGGATTTGGGGATTTGGCCCTATTGGTTAAAAAATG 60
Db 1637 CGGTCATCTTTTGATTATTAAGGATTTGGGGATTTGGCCCTATTGGTTAAAAAATG 1696
QY 61 AGCTGATTTACAAAAATTTAACGCGAATTAATCTGTGGA 101
Db 1697 AGCTGATTTACAAAAATTTAACGCGAATTAATCTGTGGA 1737
RESULT 8
LOCUS BD195386 5446 bp DNA linear PAT 17-JUL-2003
DEFINITION Composition and methods for administering Pneumococcal DNA.
ACCESSION BD195386
VERSION BD195386.1 GI:33005156
KEYWORDS JP 2002514061-A/3.
SOURCE unidentified
ORGANISM unidentified
REFERENCE
    1 (bases 1 to 5446)
    Briles, D.E., McDaniel, L.S. and Curriel, D.T.
    Composition and methods for administering Pneumococcal DNA
    Patent: JP 2002514061-A 3 14-MAY-2002;
    UNIVERSITY OF ALABAMA AT BIRMINGHAM
    OS Unidentified
    PN JP 2002514061-A/3
    PD 14-MAY-2002
    PF 04-DEC-1997 JP 1998525895
    PR 04-DEC-1996 US 08/759505
    PI DAVID E BRILES, LARRY S MCDANIEL, DAVID T CURIEL PC
    CC C12P21/06, C12N15/00, C07H21/02, C07H21/04
    CC Strandedness: Single;
    CC Topology: Linear;
    CC Composition and methods for administering Pneumococcal DNA FH
    FT Key Location/Qualifiers
    FT source 1..5446
    FT /organism='Unidentified'.
FEATURES
    source
        1..5446
        /organism='unidentified'
        /mol_type='genomic DNA'
        /db_xref='taxon:32644'
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 5446;
Best Local Similarity 100.0%; Pred. No. 4.7e-15;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGGTCATCTTTTGATTATTAAGGATTTGGGGATTTGGCCCTATTGGTTAAAAAATG 60
Db 1637 CGGTCATCTTTTGATTATTAAGGATTTGGGGATTTGGCCCTATTGGTTAAAAAATG 1696
QY 61 AGCTGATTTACAAAAATTTAACGCGAATTAATCTGTGGA 101
Db 1697 AGCTGATTTACAAAAATTTAACGCGAATTAATCTGTGGA 1737
```

```
Best Local Similarity 100.0%; Pred. No. 4.7e-15;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGTCATCTTTTGGATTATAAGGATTTTGGGATTCGGCTATTGGTTAAAAAATG 60
Db 1651 CGGTCATCTTTTGGATTATAAGGATTTTGGGATTCGGCTATTGGTTAAAAAATG 1710

Qy 61 AGCTGATTTTACAAAATTTAACCGGAATTAATCTGTGGA 101
Db 1711 AGCTGATTTTACAAAATTTAACCGGAATTAATCTGTGGA 1751

RESULT 9
AX319694 LOCUS 5446 bp DNA linear PAT 14-DEC-2001
DEFINITION Sequence 5 from Patent WO0181614.
ACCESSION AX319694
VERSION AX319694.1 GI:17901350
KEYWORDS synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Leng,J.
TITLE Cell proliferation assay
JOURNAL Patent: WO 0181614-A 5 01-NOV-2001;
Chemicon International (US)
FEATURES
source
1. 5446
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="pcDNA3 vector sequence"

ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 5446;
Best Local Similarity 100.0%; Pred. No. 4.7e-15;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGTCATCTTTTGGATTATAAGGATTTTGGGATTCGGCTATTGGTTAAAAAATG 60
Db 1651 CGGTCATCTTTTGGATTATAAGGATTTTGGGATTCGGCTATTGGTTAAAAAATG 1710

Qy 61 AGCTGATTTTACAAAATTTAACCGGAATTAATCTGTGGA 101
Db 1711 AGCTGATTTTACAAAATTTAACCGGAATTAATCTGTGGA 1751

RESULT 10
AY437643 LOCUS 5639 bp DNA circular SYN 10-NOV-2003
DEFINITION Expression vector pcGlobin 2, complete sequence.
ACCESSION AY437643
VERSION AY437643.1 GI:38155839
KEYWORDS Expression vector pcGlobin 2
ORGANISM Expression vector pcGlobin 2
REFERENCE 2
AUTHORS Ro.H., Kim,E.J. and Rhee,M.
TITLE A new vector system, pcGlobin 2 for in vitro synthesized RNA
injection into zebrafish embryos
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 5639)
AUTHORS Ro.H., Kim,E.J. and Rhee,M.
TITLE Direct Submision
JOURNAL Submitted (14-OCT-2003) Department of Biology, College of Natural
Sciences, Chungnam National University, 305-764, Daejeon 305-764,
Korea
FEATURES
source
1. 5639
/organism="Expression vector pcGlobin 2"
/mol_type="other DNA"
```

```
/db_xref="taxon:254096"
/note="eukaryotic expression vector for zebrafish embryo
microinjection; derivative of pcDNA3"
complement(4643..5503)
/codon_start=1
/product="beta-lactamase"
/protein_id="AAR12689.1"
/db_xref="GI:38155840"
/translation="MSIQHFRVALIIPFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
IEDLNSGKILESFRPEERPMWTSFKVLICGAVLSRIDAGQEQLRRIRHYSQNDLVE
YSPVTEKHLTDGMTVRELCSAAITMSDNTANALLLTIGGPKELTAFIHNNGDHVTLL
DRWEPELNEAIPNDERDTTPVAMATLRKLITGLLTLLASRQQLIDWMEADKVGPL
LRSLPAGWFIADKSGAGERSGIILAAALGPDGKPSRIVVIYITGSGQATWMDERNRQIA
EIGASLIKHW"

ORIGIN
Query Match 100.0%; Score 101; DB 12; Length 5639;
Best Local Similarity 100.0%; Pred. No. 4.7e-15;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGTCATCTTTTGGATTATAAGGATTTTGGGATTCGGCTATTGGTTAAAAAATG 60
Db 1844 CGGTCATCTTTTGGATTATAAGGATTTTGGGATTCGGCTATTGGTTAAAAAATG 1903

Qy 61 AGCTGATTTTACAAAATTTAACCGGAATTAATCTGTGGA 101
Db 1904 AGCTGATTTTACAAAATTTAACCGGAATTAATCTGTGGA 1944

RESULT 11
AX211282 LOCUS 5651 bp DNA linear PAT 06-SEP-2001
DEFINITION Sequence 6 from Patent WO0158493.
ACCESSION AX211282
VERSION AX211282.1 GI:15523691
KEYWORDS synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Schambye,H.T., Andersen,K.V., van den Hazel,B., Christiansen,J. and
Jeppesen,C.B.
TITLE Conjugates of follicle stimulating hormones
JOURNAL Patent: WO 0158493-A 6 16-AUG-2001;
Maxygen Aps (DK)
FEATURES
source
1. 5651
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Coding sequence for human FSH-beta"

ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 5651;
Best Local Similarity 100.0%; Pred. No. 4.7e-15;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGTCATCTTTTGGATTATAAGGATTTTGGGATTCGGCTATTGGTTAAAAAATG 60
Db 2273 CGGTCATCTTTTGGATTATAAGGATTTTGGGATTCGGCTATTGGTTAAAAAATG 2332

Qy 61 AGCTGATTTTACAAAATTTAACCGGAATTAATCTGTGGA 101
Db 2333 AGCTGATTTTACAAAATTTAACCGGAATTAATCTGTGGA 2373

RESULT 12
AX349366 LOCUS 5651 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 4 from Patent WO0202597.
ACCESSION AX349366
VERSION AX349366.1 GI:18615329
```

KEYWORDS	synthetic construct
SOURCE	synthetic construct
ORGANISM	other sequences; artificial sequences.
REFERENCE	1
AUTHORS	Okkels,J.S., Jensen,A.D. and van den Hazel,B.C.
TITLE	Peptide extended glycosylated polypeptides
JOURNAL	Patent: WO 0202597-A 4 10-JAN-2002;
JOURNAL	Maxygen Aps (DK) ; Maxygen Holdings Ltd (KY)
FEATURES	Location/Qualifiers
source	1. .5651
	/organism="synthetic construct"
	/mol_type="unassigned DNA"
exon	1231. .1617
	/db_xref="taxon:32630"
	/note="Coding sequence for human FSH-beta"
ORIGIN	
Query Match	100.0%; Score 101; DB 6; Length 5651;
Best Local Similarity	100.0%; Pred. No. 4.7e-15;
Matches 101; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 CGGTCCTATCTTTGATTTTAAAGGATTTTGGGATTTTCGCCCTATTGTTTAAAAAATG 60
Db	2273 CGGTCCTATCTTTGATTTTAAAGGATTTTGGGATTTTCGCCCTATTGTTTAAAAAATG 2332
QY	61 AGCTGATTTTACAAAAATTTAACGGCAATTAATTCGTGGA 101
Db	2333 AGCTGATTTTACAAAAATTTAACGGCAATTAATTCGTGGA 2373
RESULT 13	
AX202478	
LOCUS	AX202478 5731 bp DNA linear PAT 30-AUG-2001
DEFINITION	Sequence 66 from Patent WO0152620.
ACCESSION	AX202478
VERSION	AX202478.1 GI:15392206
KEYWORDS	synthetic construct
SOURCE	synthetic construct
ORGANISM	other sequences; artificial sequences.
REFERENCE	1
AUTHORS	Barbas,C.F., Stege,J.T., Guan,X. and Dalmia,B.
TITLE	Methods and compositions to modulate expression in plants
JOURNAL	Patent: WO 0152620-A 66 26-JUL-2001;
JOURNAL	The Scripps Research Institute (US) ; SYNGENTA AGRICULTURAL
JOURNAL	DISCOVERY, INC. (CA)
FEATURES	Location/Qualifiers
source	1. .5731
	/organism="synthetic construct"
	/mol_type="unassigned DNA"
	/db_xref="taxon:32630"
	/note="2C7-SID"
ORIGIN	
Query Match	100.0%; Score 101; DB 6; Length 5731;
Best Local Similarity	100.0%; Pred. No. 4.7e-15;
Matches 101; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 CGGTCCTATCTTTGATTTTAAAGGATTTTGGGATTTTCGCCCTATTGTTTAAAAAATG 60
Db	2353 CGGTCCTATCTTTGATTTTAAAGGATTTTGGGATTTTCGCCCTATTGTTTAAAAAATG 2412
QY	61 AGCTGATTTTACAAAAATTTAACGGCAATTAATTCGTGGA 101
Db	2413 AGCTGATTTTACAAAAATTTAACGGCAATTAATTCGTGGA 2453
RESULT 14	
AX685746	
LOCUS	AX685746 5995 bp DNA linear PAT 29-MAR-2003
DEFINITION	Sequence 5 from Patent WO02102854.
ACCESSION	AX685746

VERSION	AX685746.1 GI:29371751
KEYWORDS	synthetic construct
SOURCE	synthetic construct
ORGANISM	other sequences; artificial sequences.
REFERENCE	1
AUTHORS	Thomassen-Wolf,E., Borges,E., Yayon,A. and Rom,E.
TITLE	Antibodies that block receptor protein tyrosine kinase activation,
JOURNAL	Methods of screening for and uses thereof
JOURNAL	Patent: WO 02102854-A 5 27-DEC-2002;
JOURNAL	MorphoSys AG (DE) ; ProChon Biotech Ltd. (IL)
FEATURES	Location/Qualifiers
source	1. .5995
	/organism="synthetic construct"
	/mol_type="unassigned DNA"
	/db_xref="taxon:32630"
	<963. .1670
	/note="unnamed protein product"
	/codon_start=1
	/transl_table=11
	/protein_id="CAD86574.1"
	/db_xref="GI:29371752"
	/translation="DPEPKSCDKHTKCPPELGGPSVFLPFPKPKDTLMISRT
	PEVTCVVDVSHEDPEVKFNWYDGEVHNNAKTRPEEQYNSTRVSVLTVLHQDWL
	NGKEYCKVKSNKALPAPIETKTSKAGQPEQYVTLPPSRDELTKNQVSLTCLVKGF
	YPSDIAVWESNGQPENNYKTTPPVLDSGSEFFLYSKLTVDKSRWQQGNVFPSCSYMHE
	ALHNHYTKSLSLSPGK"
ORIGIN	
Query Match	100.0%; Score 101; DB 6; Length 5995;
Best Local Similarity	100.0%; Pred. No. 4.6e-15;
Matches 101; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 CGGTCCTATCTTTGATTTTAAAGGATTTTGGGATTTTCGCCCTATTGTTTAAAAAATG 60
Db	2497 CGGTCCTATCTTTGATTTTAAAGGATTTTGGGATTTTCGCCCTATTGTTTAAAAAATG 2556
QY	61 AGCTGATTTTACAAAAATTTAACCGCAATTAATTCGTGGA 101
Db	2557 AGCTGATTTTACAAAAATTTAACCGCAATTAATTCGTGGA 2597
RESULT 15	
GGAS75208	
LOCUS	GGAS75208 6084 bp DNA circular SYN 03-JUL-2003
DEFINITION	Expression vector pCLGPPA.
ACCESSION	AJ575208
VERSION	AJ575208.1 GI:32451228
KEYWORDS	Expression vector pCLGPPA
SOURCE	Expression vector pCLGPPA
ORGANISM	other sequences; artificial sequences; vectors.
REFERENCE	1
AUTHORS	Scaal,M., Gros,J., Lesbros,C. and Marcelle,C.
TITLE	In ovo electroporation of avian somites
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 6084)
AUTHORS	Marcelle,C.
TITLE	Direct Submision
JOURNAL	Submitted (13-JUN-2003) Marcelle C., Lgpd, Institut de Biologie du
JOURNAL	Developpement, Campus de Luminy Case 907 F-13288 Marseille, 13288,
JOURNAL	FRANCE
FEATURES	Location/Qualifiers
source	1. .6084
	/organism="Expression vector pCLGPPA"
	/mol_type="other DNA"
	/db_xref="taxon:236984"
	/focus
	12. .397
	/organism="unidentified cytomegalovirus"
	/mol_type="other DNA"
	/db_xref="taxon:205912"
	398. .1652
source	

```

/organism="Gallus gallus"
/mol_type="other DNA"
/db_xref="taxon:9031"
1653..1733
/organism="Oryctolagus cuniculus"
/mol_type="other DNA"
/db_xref="taxon:9986"
1782..1996
/organism="Bos taurus"
/mol_type="other DNA"
/db_xref="taxon:9913"
2442..2836
/organism="Simian virus 40"
/mol_type="other DNA"
/db_xref="taxon:10633"
2908..3723
/organism="Aequorea victoria"
/mol_type="other DNA"
/db_xref="taxon:6100"
12..399
740..887
/gene="beta actin"
740..887
/gene="beta actin"
1734..1776
/note="multiple cloning site; MCS
under control of CMV enhancer/chicken beta-actin promoter"
1782..1996
/note="bovine growth hormone"
2442..2836
2908..3723
/note="eGFP
under control of SV40 promoter"
3757..3887
/note="late"

```

ORIGIN

```

Query Match      100.0%; Score 101; DB 12; Length 6084;
Best Local Similarity 100.0%; Pred. No. 4.6e-15;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  CGGTCATTCTTTTGAATTATAGGGATTTCGGGATTTCGGCCTATTGGTTAAAAAATG 60
      |||
Db      2398  CGGTCATTCTTTTGAATTATAGGGATTTCGGGATTTCGGCCTATTGGTTAAAAAATG 2457

QY      61  AGCTGATTTAACAAAAATTAAACGCGAATTAATCTGTGGA 101
      |||
Db      2458  AGCTGATTTAACAAAAATTAAACGCGAATTAATCTGTGGA 2498

```

Search completed: July 14, 2005, 14:03:21
Job time : 757.618 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:35:42 ; Search time 142.398 Seconds
(without alignments)
4198.742 Million cell updates/sec

Title: US-09-482-682-8_COPY_3674_3774
Perfect score: 101
Sequence: 1 cggctattctttgattta.....acgcgaattaattctgtgga 101

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	3482	2	ADH11353
2	101	100.0	4597	4	AAT24901
3	101	100.0	4639	6	AAD339652
4	101	100.0	5015	10	ADB33528
5	101	100.0	5070	4	AAS12839
6	101	100.0	5082	2	ADH11417
7	101	100.0	5173	6	ABK88869
8	101	100.0	5173	12	ADE83792
9	101	100.0	5173	12	ADO06721
10	101	100.0	5218	12	ADM97811
11	101	100.0	5302	12	ADI34681
12	101	100.0	5425	2	ADH11233
13	101	100.0	5431	6	ABN86685
14	101	100.0	5431	10	ADE21866
15	101	100.0	5431	12	ADO05277
16	101	100.0	5432	3	AAZ28947
17	101	100.0	5446	2	AAV38297
18	101	100.0	5446	6	AAI18619
19	101	100.0	5446	6	ABL53540
20	101	100.0	5446	12	ADN36314

21	101	100.0	5458	6	ABL58494
22	101	100.0	5458	6	ABL58493
23	101	100.0	5543	6	ABK88868
24	101	100.0	5543	12	ADE83791
25	101	100.0	5543	12	ADO06720
26	101	100.0	5614	6	ABL58489
27	101	100.0	5614	6	ABL58490
28	101	100.0	5651	5	AAI166195
29	101	100.0	5651	6	ABK40237
30	101	100.0	5695	6	ABL58492
31	101	100.0	5695	6	ABL58491
32	101	100.0	5695	8	ABT40262
33	101	100.0	5695	8	ADA89054
34	101	100.0	5695	10	ADG74306
35	101	100.0	5731	4	AAI11615
36	101	100.0	5864	6	AAI44423
37	101	100.0	5864	6	AAI44424
38	101	100.0	6082	8	AAI56212
39	101	100.0	6082	8	AAI56211
40	101	100.0	6082	8	AAI56210
41	101	100.0	6085	8	AAI56213
42	101	100.0	6094	8	AAI56215
43	101	100.0	6097	8	AAI56214
44	101	100.0	6100	6	ABK96469
45	101	100.0	6135	6	ABK96470

ALIGNMENTS

RESULT 1
ADH11353
ID ADH11353 standard; DNA; 3482 BP.
XX
AC ADH11353;
XX
DT 11-MAR-2004 (first entry)
XX
DE Vertebrate UNC-53 protein homologue related nucleotide sequence.
XX
KW UNC-53 vertebrate protein homologue; UNC-53; Caenorhabditis elegans;
KW cell shape regulator; cell motility regulator; cell migration;
KW cell behaviour regulator; phenotype; signal transduction pathway;
KW signal transducing protein; phenotypic; signal integrator protein;
KW neuronal regeneration; revascularisation; wound healing;
KW chronic neurodegenerative disease; acute traumatic injury;
KW fibrotic disease; gene; ds.
XX
OS Unidentified.
XX
PN WO9824810-A2.
XX
PD 11-JUN-1998.
XX
PF 03-DEC-1997; 97WO-EP006956.
XX
PR 04-DEC-1996; 96GB-00025283.
XX
PA (JANC) JANSSEN PHARM NV.
XX
PI Platteauw CJ, Buesa Arjol CM, Deraeymaeker M, Verhaesselt P;
PI Pujol NJR, Maertens LJS, Luyten W, Geerts H, Vandekerckhove JS;
PI Geyse J, Bogaert TAOE,
XX
WPI; 1998-362411/31.
DR P-PSDB; ADH11354.
XX
PT Vertebrate homologues of C. elegans UNC-53 protein - useful for, e.g.
PT promoting neuronal regeneration, treating chronic neuro-degenerative
PT diseases or acute traumatic injuries.
XX
PS Disclosure; Page 414-417; 479pp; English.
XX

CC The present invention describes a vertebrate protein homologue of an UNC-53 protein of *Caenorhabditis elegans* or a functional equivalent, CC derivative or bioprecursor of UNC-53. Also described: (1) a cDNA sequence encoding a vertebrate homologue of the *C. elegans* UNC-53 protein; (2) a nucleic acid which hybridises to the cDNA of (1); (3) vector comprising the cDNA as in (1); (4) a host cell containing the vector as in (3); (5) a transgenic cell, tissue or animal comprising the vector as in (3); (6) a compound identified as an enhancer or inhibitor of the regulation of cell shape, motility, or the direction of cell migration for use as a therapeutic; (7) a method for determination of whether a protein is an inhibitor or enhancer of regulation of cell behaviour, growth, shape or motility or the direction of migration by contacting a host cell CC expressing a homologue of UNC-53 and determining a change of phenotype; CC (8) a method for identification of vertebrate homologues of *C. elegans* unc-53 gene by hybridising a probe of 15-50 bp of the unc-53 sequence to a DNA library; and (9) a method for identification of a protein which is active in the signal transduction pathway of a cell of which a vertebrate CC homologue of UNC-53 is a component comprising: (i) contacting an extract of a cell with an antibody to the UNC-53 homologue; (ii) identifying an CC antibody/homologue complex; and (iii) analysing such a complex to CC identify any non-antibody protein bound to the complex. UNC-53 is a CC signal transducing or signal integrator protein involved in controlling CC directionality of cell migration and cell shape in *C. elegans*. Vertebrate CC homologues of UNC-53 can be used to promote neuronal regeneration, CC revascularisation or wound healing, to treat chronic neurodegenerative CC diseases or acute traumatic injuries or fibrotic diseases. The present CC sequence is used in the exemplification of the present invention.

XX Sequence 3482 BP; 767 A; 956 C; 913 G; 846 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 101; DB 2; Length 3482;
Best Local Similarity 100.0%; Pred. No. 4.1e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGGCTATTCTTTTGATTATTAAGGGAATTTGGGATTTGGCCCTATTGGTTAAAAAATG 60
Db 1721 CGGCTATTCTTTTGATTATTAAGGGAATTTGGGATTTGGCCCTATTGGTTAAAAAATG 1780
QY 61 AGCTGATTTAACAAAAATTAACGCGAATTAATTCGTGGA 101
Db 1781 AGCTGATTTAACAAAAATTAACGCGAATTAATTCGTGGA 1821

RESULT 2
AAF24901
ID AAF24901 standard; DNA; 4597 BP.
XX
AC AAF24901;
XX
DT 20-APR-2001 (first entry)
XX
DE Nucleotide sequence of the plasmid pCDNA3.1/GS.
XX
KW Microsphere; dihydrazide; hyaluronic acid; inflammatory response;
KW myocardial ischemia; cardiac angiogenesis; haemophilia;
KW vascular endothelial growth factor; VEGF; ss.
XX
OS Synthetic.
XX
PN WO200078358-A2.
XX
PD 28-DEC-2000.
XX
PF 19-JUN-2000; 2000WO-US016837.
XX
PR 18-JUN-1999; 99US-0140260P.
XX
PA (COLL-) COLLABORATIVE GROUP LTD.
XX
PI Chen W;
XX
DR WPI; 2001-071363/08.
XX

PT Hyaluronic acid micro spheres for use in gene therapy of myocardial
PT ischemia and hemophilia, comprising dihydrazide derivatized hyaluronic
PT acids crosslinked to nucleic acids.
XX
PS Example 1; Page 36-38; 38pp; English.
XX
CC The specification describes a microsphere comprising dihydrazide
CC derivatised hyaluronic acid crosslinked to a nucleic acid (NA). The
CC microspheres cause reduced inflammatory responses, and have increased
CC safety and biodegradability. The microspheres are useful for transfecting
CC a cell of a subject and for treating a subject having myocardial
CC ischemia, by increasing cardiac angiogenesis. They are also useful for
CC treating haemophilia. The present sequence represents the plasmid
CC pCDNA3.1/GS, into which is inserted a polynucleotide sequence which is
CC crosslinked to hyaluronic acid. The polynucleotide sequence encodes a
CC vascular endothelial growth factor (VEGF)
XX
SQ Sequence 4597 BP; 1062 A; 1214 C; 1206 G; 1115 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 4; Length 4597;
Best Local Similarity 100.0%; Pred. No. 4.3e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGGCTATTCTTTTGATTATTAAGGGAATTTGGGATTTGGCCCTATTGGTTAAAAAATG 60
Db 2227 CGGCTATTCTTTTGATTATTAAGGGAATTTGGGATTTGGCCCTATTGGTTAAAAAATG 2286
QY 61 AGCTGATTTAACAAAAATTAACGCGAATTAATTCGTGGA 101
Db 2287 AGCTGATTTAACAAAAATTAACGCGAATTAATTCGTGGA 2327

RESULT 3
AAD39652
ID AAD39652 standard; DNA; 4639 BP.
XX
AC AAD39652;
XX
DT 22-OCT-2002 (first entry)
XX
DE Human small nuclear RNA (snRNA) DNA.
XX
KW Human; recombinant vector; insertion cassette; small nuclear RNA; snRNA;
KW transgenic animal; ds.
XX
OS Homo sapiens.
XX
PN US2002058287-A1.
XX
PD 16-MAY-2002.
XX
PF 12-MAR-2001; 2001US-00804481.
XX
PR 10-MAR-2000; 2000US-0188304P.
XX
PA (WHEED) WHITEHEAD INST BIOMEDICAL RES.
XX
PI Graaf DD, Lander ES;
XX
XX WPI; 2002-499510/53.
XX
DR New recombinant vector containing sequence for small nuclear RNA, useful
PT e.g. for identifying variant snRNA that suppresses expression of
PT transcription products.
XX
PS Disclosure; Fig 1; 18pp; English.
XX
CC The invention relates to a recombinant vector which comprises DNA,
CC consisting of an insertion cassette contained between at least two
CC insertion sites, that encodes a small nuclear (sn) RNA. The invention is
CC used to identify snRNA modifications that inhibit expression of
CC transcription products (and the identified snRNA are used to suppress
CC expression) for delivering antisense sequences to the nucleus and to

CC create transgenic animals. The present DNA sequence is human snRNA, UI
SQ Sequence 4639 BP; 1067 A; 1198 C; 1243 G; 1131 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 6; Length 4639;
Best Local Similarity 100.0%; Pred. No. 4.3e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGGTCTATCTTTGATTATTAAGGATTTTCGGGATTTTCGCCCTATTGGTTAAAAAATG 60
DB 1261 CGGTCTATCTTTGATTATTAAGGATTTTCGGGATTTTCGCCCTATTGGTTAAAAAATG 1320
QY 61 AGCTGATTTTAAACAAAATTTAAGCGAATTAATCTGTGGA 101
DB 1321 AGCTGATTTTAAACAAAATTTAAGCGAATTAATCTGTGGA 1361
RESULT 4
ID ADB33528 standard; DNA; 5015 BP.
XX AC ADB33528;
XX DT 04-DEC-2003 (first entry)
XX DE Expression vector nucleotide sequence SEQ ID NO:27.
XX KW fusion protein; amyloid precursor protein; APP; transcription factor;
KW neurotrophic; neuroprotective; APP inhibitor;
KW amyloid precursor protein inhibitor; Alzheimer's disease; beta-secretase;
KW gamma-secretase; human; gene; ds.
XX OS Synthetic.
XX OS Homo sapiens.
XX PN WO2003072041-A2.
XX PD 04-SEP-2003.
XX PF 23-FEB-2003; 2003WO-US005458.
XX PR 27-FEB-2002; 2002US-0360274P.
XX PA (MERI) MERCK & CO INC.
XX PI Espeseth AS, Ferrer M, Flores OA, Hazuda DJ, Inglese J;
PI Miller MD, Register B, Shi X, Simon AJ, Zuck PD;
XX WPI; 2003-689968/65.
XX DNA encoding a fusion protein of amyloid precursor protein, useful in
PT screening for anti-Alzheimer agents, comprises a fused transcription
PT factor.
XX PS Disclosure; Fig 32B-F; 193pp; English.
XX The present invention describes a DNA molecule (I) that encodes a fusion
CC protein (FP) comprising: (i) an amino acid sequence of amyloid precursor
CC protein (APP), either the wild type, Swedish or NREV versions; and (ii) a
CC transcription factor (TF), fused in frame to the C-terminus of (i). Also
CC described: (1) an expression vector containing (I); (2) a eukaryotic cell
CC containing (I); and (3) methods for identifying a compound (A) that
CC inhibits processing of APP, using the cells of (2). (I) has neurotropic and
CC neuroprotective activities. (I) can be used to produce eukaryotic cells
CC that express FP and are useful in screening for agents that inhibit
CC processing of APP. The agents are potentially useful for the treatment or
CC prevention of Alzheimer's disease. Cells that express FP can screen for
CC inhibitors of: (a) beta- and gamma-secretases; and (b)
CC cytoplasmic/extracellular APP signaling in a single assay. Cell-based
CC assays may be free of interference from alpha-secretase activity and are
CC homogeneous (no chromatography, immunoprecipitation or washing required)
CC so well suited to high-throughput screening. The present sequence
CC represents a plasmid nucleotide sequence from the present invention.

XX SQ Sequence 5015 BP; 1167 A; 1297 C; 1279 G; 1272 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 10; Length 5015;
Best Local Similarity 100.0%; Pred. No. 4.3e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGGTCTATCTTTGATTATTAAGGATTTTCGGGATTTTCGCCCTATTGGTTAAAAAATG 60
DB 1637 CGGTCTATCTTTGATTATTAAGGATTTTCGGGATTTTCGCCCTATTGGTTAAAAAATG 1696
QY 61 AGCTGATTTTAAACAAAATTTAAGCGAATTAATCTGTGGA 101
DB 1697 AGCTGATTTTAAACAAAATTTAAGCGAATTAATCTGTGGA 1737
RESULT 5
ID AAS12839 standard; DNA; 5070 BP.
XX AC AAS12839;
XX DT 21-NOV-2001 (first entry)
XX DE DNA sequence of pSecTag1 vector used to express VEGF-A/VEGF-C hybrids.
XX KW Human; vascular endothelial growth factor; VEGF-A; vasculogenesis;
KW angiogenesis; blood vessel; cancer; proliferative retinopathy; psoriasis;
KW age-related macular degeneration; rheumatoid arthritis; cardiovascular;
KW pSecTag1; cyclic; circular; ds.
XX OS Synthetic.
XX OS WO200162942-A2.
XX PN 30-AUG-2001.
XX PD 26-FEB-2001; 2001WO-US006113.
XX PF 25-FEB-2000; 2000US-0185205P.
XX PR 18-MAY-2000; 2000US-0205331P.
XX PA (LUDW-) LUDWIG INST CANCER RES.
XX PA (LICN) LICENTIA OY.
XX PI Alitalo K, Jeltsch MM;
XX WPI; 2001-536640/59.
XX PT Polypeptides that bind cellular receptors for vascular endothelial growth
PT factors, polynucleotides encoding them.
XX PS Example 1; Page 173-176; 261pp; English.
XX The present invention relates to polypeptides that bind cellular
CC receptors for vascular endothelial growth factors (VEGFs), the
CC polynucleotides encoding them, and their use for identifying agents that
CC modulate interactions between VEGFs and their receptors. VEGFs and their
CC receptors play an important role in vasculogenesis, the development of
CC the embryonic vasculature from early differentiating endothelial cells
CC and angiogenesis, the process of forming new blood vessels from pre-
CC existing ones. Modulators of interactions between VEGF and its receptors
CC may be used to treat dysfunction of the endothelial cell regulatory
CC system. Such disorders include cancers, abnormal angiogenesis,
CC proliferative retinopathies, age-related macular degeneration, rheumatoid
CC arthritis and psoriasis. The polypeptides of the invention exhibit unique
CC receptor binding profiles compared to known naturally occurring VEGFs.
CC The present DNA sequence for pSecTag1 vector is used to express VEGF-
CC A/VEGF-C hybrids in the methods of the present invention
XX SQ Sequence 5070 BP; 1186 A; 1308 C; 1288 G; 1288 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 4; Length 5070;

Best Local Similarity 100.0%; Pred. No. 4.3e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGCTATCTTTGATTATTAAGGATTTGGGGATTTCCGCCCTATTGGTTAAAAAATG 60
|||||
Db 1693 CGGCTATCTTTGATTATTAAGGATTTGGGGATTTCCGCCCTATTGGTTAAAAAATG 1752

QY 61 AGCTGATTTACAAAAATTAACGCGAATTAATCTCTGTGA 101
|||||
Db 1753 AGCTGATTTACAAAAATTAACGCGAATTAATCTCTGTGA 1793

RESULT 6
ADH11417
ID ADH11417 standard; DNA; 5082 BP.
AC ADH11417;
XX
DT 11-MAR-2004 (first entry)
XX
DE Plasmid pCB201 nucleotide sequence.
XX
KW UNC-53 vertebrate protein homologue; UNC-53; Caenorhabditis elegans;
KW cell shape regulator; cell motility regulator; cell migration;
KW cell behaviour regulator; phenotype; signal transduction pathway;
KW signal transducing protein; signal integrator protein;
KW neuronal regeneration; revascularisation; wound healing;
KW chronic neurodegenerative disease; acute traumatic injury;
KW fibrotic disease; human; gene; ds.
XX
OS Synthetic.
OS Homo sapiens.
XX

FH Key Location/Qualifiers
FT CDS 1028..2149
FT /*tag= a
FT
XX
FN WO9824810-A2.
XX
PD 11-JUN-1998.
XX
PF 03-DEC-1997; 97WO-EP006956.
XX
PR 04-DEC-1996; 96GB-00025283.
XX
XX (JANC) JANSSEN PHARM NV.
XX
XX Platreeuw CJ, Buesa Arjol CM, Deraeymaeker M, Verhasselt P;
PI Pujol NJR, Maertens LJS, Luyten W, Geerts H, Vandekerckhove JS;
PI Geysen J, Bogaert TAOE;
XX
XX WPI; 1998-362411/31.
DR P-PSDB; ADH11424.
XX
XX Vertebrate homologues of C. elegans UNC-53 protein - useful for, e.g.
PT promoting neuronal regeneration, treating chronic neuro-degenerative
PT diseases or acute traumatic injuries.
XX
XX Claim 96; SEQ ID NO 10; 479pp; English.
XX
XX The present invention describes a vertebrate protein homologue of an UNC-
CC 53 protein of Caenorhabditis elegans or a functional equivalent,
CC derivative or bioprecursor of UNC-53. Also described: (1) a cDNA sequence
CC encoding a vertebrate homologue of the C. elegans UNC-53 protein; (2) a
CC nucleic acid which hybridises to the cDNA of (1); (3) vector comprising
CC the cDNA as in (1); (4) a host cell containing the vector as in (3); (5)
CC a transgenic cell, tissue or animal comprising the vector as in (3); (6)
CC a compound identified as an enhancer or inhibitor of the regulation of
CC cell shape, motility, or the direction of cell migration for use as a
CC therapeutic; (7) a method for determination of whether a protein is an
CC inhibitor or enhancer of regulation of cell behaviour, growth, shape or
CC motility or the direction of migration by contacting a host cell
CC expressing a homologue of UNC-53 and determining a change of phenotype;

(8) a method for identification of vertebrate homologues of C. elegans
unc-53 gene by hybridising a probe of 15-50 bp of the unc-53 sequence to
a DNA library; and (9) a method for identification of a protein which is
active in the signal transduction pathway of a cell of which a vertebrate
homologue of UNC-53 is a component comprising: (i) contacting an extract
of a cell with an antibody to the UNC-53 homologue; (ii) identifying an
antibody/homologue complex; and (iii) analysing such a complex to
identify any non-antibody protein bound to the complex. UNC-53 is a
signal transducing or signal integrator protein involved in controlling
directionality of cell migration and cell shape in C. elegans. Vertebrate
homologues of UNC-53 can be used to promote neuronal regeneration,
revascularisation or wound healing, to treat chronic neurodegenerative
diseases or acute traumatic injuries or fibrotic diseases. The present
sequence is used in the exemplification of the present invention.

XX
SQ Sequence 5082 BP; 1164 A; 1365 C; 1311 G; 1242 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 2; Length 5082;
Best Local Similarity 100.0%; Pred. No. 4.3e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGCTATCTTTGATTATTAAGGATTTGGGGATTTCCGCCCTATTGGTTAAAAAATG 60
|||||
Db 3321 CGGCTATCTTTGATTATTAAGGATTTGGGGATTTCCGCCCTATTGGTTAAAAAATG 3380

QY 61 AGCTGATTTACAAAAATTAACGCGAATTAATCTCTGTGA 101
|||||
Db 3381 AGCTGATTTACAAAAATTAACGCGAATTAATCTCTGTGA 3421

RESULT 7
ABK88869
ID ABK88869 standard; DNA; 5173 BP.
XX
AC ABK88869;
XX
DT 07-AUG-2003 (revised)
DT 07-OCT-2002 (first entry)
XX
DE Topoisomerase vector pcDNA6.2/V5/GWD-TOPO.
XX
KW ds; topoisomerase recognition site; topoisomerase; pcDNA6WDT(sc);
KW PENTR-DT(sc); pcDNA-DEST41; PENTR/D-TOPO; PENTR/SD/D-TOPO;
KW pcDNA3.2/V5/GWD-TOPO; pcDNA6.2/V5/GWD-TOPO; recombinational cloning;
KW gene targeting; mutation; cyclic; circular; vector.
XX
OS Escherichia coli.
OS Viruses.
OS Human cytomegalovirus.
OS Synthetic.
XX
XX WO200246372-A1.
XX
XX 13-JUN-2002.
XX
XX 07-DEC-2001; 2001WO-US045773.
XX
XX 08-DEC-2000; 2000US-0254510P.
PR 11-DEC-2000; 2000US-00732914.
PR 14-SEP-2001; 2001US-0318902P.
PR 28-SEP-2001; 2001US-0326092P.
PR 27-NOV-2001; 2001US-0333124P.
XX
XX (INVI-) INVITROGEN CORP.
XX
XX Chesnut JD, Carrino J, Leong L, Madden K, Gleeson M, Fan J;
PI Brasch MA, Cheo D, Hartley JL, Byrd DRN, Temple GF;
XX
XX WPI; 2002-519662/55.
XX
XX New isolated nucleic acid molecule comprises one or more recombination
PT sites and one or more topoisomerase recognition sites and/or one or more
PT topoisomerases, useful in recombinational cloning.

XX Disclosure; Fig 25B-25C; 324pp; English.

XX The invention relates to an isolated nucleic acid molecule (1)

XX comprising: (a) one or more recombination sites; and (b) one or more

XX topoisomerase recognition sites and/or one or more topoisomerases. Also

XX included are a vector comprising the nucleic acid, a vector chosen from

XX pCDNAGWDT(sc), pENTR-DT(sc), pCDNA-DEST41, pENTR/D-TOPO, pENTR/SD/D-TOPO,

XX pCDNA3, 2/V5/GWD-TOPO or pCDNA6.2/V5/GWD-TOPO, a host cell comprising the

XX nucleic acid or vectors and an in vitro method of cloning a nucleic acid

XX molecule involving: (a) obtaining a first nucleic acid molecule to be

XX cloned; (b) mixing the first nucleic acid molecule to be cloned in vitro

XX with a second nucleic acid molecule comprising at least a first

XX topoisomerase recognition site flanked by at least a first recombination

XX site, and at least a second topoisomerase recognition site flanked by at

XX least a second recombination site, where the first and second

XX recombination sites do not recombine with each other, and at least one

XX topoisomerase; and (c) incubating the mixture under conditions such that

XX the first nucleic acid molecule to be cloned is inserted into the second

XX nucleic acid molecule between the first and second topoisomerase

XX recognition sites, thereby producing a first product molecule comprising

XX the first nucleic acid molecule to be cloned between the first and second

XX recombination sites. The method is useful for cloning a nucleic acid

XX molecule. The nucleic acid (1) is useful in methods for recombinational

XX cloning and facilitates construction of gene targeting nucleic acid

XX molecules or vectors which may be used to knockout or mutate a sequence

XX or gene of interest, particularly genes or sequences within a host or

XX host cells such as animal, plant, etc. Thus the nucleic acid is most

XX preferably used for targeting or mutating a sequence of gene. The present

XX sequence is the topoisomerase site-containing vector pCDNA6.2/V5/GWD-

XX TOPO. (Updated on 07-AUG-2003 to correct OS field.)

XX SQ Sequence 5173 BP; 1247 A; 1338 C; 1274 G; 1305 T; 0 U; 9 Other;

Query Match 100.0%; Score 101; DB 6; Length 5173;

Best Local Similarity 100.0%; Pred. No. 4.3e-19;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGCTATTCTTTGATTATTAAGGATTTTGGGATTTTCGGCTATTGTTAAAAAATG 60

Db 1747 CGGCTATTCTTTGATTATTAAGGATTTTGGGATTTTCGGCTATTGTTAAAAAATG 1806

QY 61 AGCTGATTTAACAAAAATTAACGCGAATTAATTCGTGGA 101

Db 1807 AGCTGATTTAACAAAAATTAACGCGAATTAATTCGTGGA 1847

RESULT 8

ADE83792

ID ADE83792 standard; DNA; 5173 BP.

XX AC ADE83792;

XX DT 29-JAN-2004 (first entry)

XX DE Plasmid pCDNA6.2/V5/GWD-TOPO.

XX KW recombination site; topoisomerase recognition site; topoisomerase;

XX KW transfection; two-hybrid assay; ds; plasmid; cyclic; circular;

XX KW pCDNA6.2/V5/GWD-TOPO.

XX OS Unidentified.

XX PN US2003186233-A1.

XX PD 02-OCT-2003.

XX PF 07-DEC-2001; 2001US-00005876.

XX PR 08-DEC-2000; 2000US-0254510P.

XX PR 11-DEC-2000; 2000US-00732914.

XX PR 21-MAY-2001; 2001US-0291972P.

XX PR 14-SEP-2001; 2001US-0318902P.

PR 28-SEP-2001; 2001US-0326092P.

PR 27-NOV-2001; 2001US-0333124P.

XX (INVI-) INVITROGEN CORP.

XX Chesnut JD, Carrino J, Leong L, Madden K, Gleeson M, Fan J;

PI Brasch MA, Cheo D, Hartley JL, Byrd DRN, Temple GF;

XX WPI; 2004-031998/03.

XX Isolated nucleic acid molecule useful as vector in host cell comprises at

XX least one recombination site, and at least one topoisomerase recognition

XX site and/or at least one topoisomerase.

XX Disclosure; Fig 25B-C; 136pp; English.

XX The invention describes an isolated nucleic acid molecule comprising at

XX least one recombination site, and at least one topoisomerase recognition

XX site and/or at least one topoisomerase. The isolated nucleic acid

XX molecule is used as a vector in a host cell. It can be used directly for

XX transfecting a cell, or as a template for performing amplification, e.g.

XX PCR, a recombination reaction, an in vitro transcription reaction, or a

XX coupled transcription/translation reaction. The invention allows several

XX nucleic acid molecules to be covalently linked in a predetermined

XX directional orientation. A functional product can be selected in vitro by

XX performing an amplification reaction using primers specific for the

XX termini of the desired covalently linked recombinant nucleic acid

XX molecule. This sequence represents plasmid pCDNA6.2/V5/GWD-TOPO

XX associated with the isolation and analysis of the novel polynucleotide of

XX the invention.

XX SQ Sequence 5173 BP; 1247 A; 1338 C; 1274 G; 1305 T; 0 U; 9 Other;

Query Match 100.0%; Score 101; DB 12; Length 5173;

Best Local Similarity 100.0%; Pred. No. 4.3e-19;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGCTATTCTTTGATTATTAAGGATTTTGGGATTTTCGGCTATTGTTAAAAAATG 60

Db 1747 CGGCTATTCTTTGATTATTAAGGATTTTGGGATTTTCGGCTATTGTTAAAAAATG 1806

QY 61 AGCTGATTTAACAAAAATTAACGCGAATTAATTCGTGGA 101

Db 1807 AGCTGATTTAACAAAAATTAACGCGAATTAATTCGTGGA 1847

RESULT 9

ADO06721

ID ADO06721 standard; DNA; 5173 BP.

XX AC ADO06721;

XX DT 01-JUL-2004 (first entry)

XX DE Recombinatorial cloning method plasmid vector - pCDNA6.2/V5/GWD-TOPO.

XX KW vector; recombination site; topoisomerase recognition site;

XX KW topoisomerase; in vitro cloning; recombinatorial cloning; plasmid;

XX KW pCDNA6.2/V5/GWD-TOPO; ds.

XX OS Unidentified.

XX PN WO2003103600-A2.

XX PD 18-DEC-2003.

XX PF 05-JUN-2003; 2003WO-US018036.

XX PR 05-JUN-2002; 2002US-0385613P.

XX PA (INVI-) INVITROGEN CORP.

XX PI Chesnut JD, Carrino J, Leong L, Madden K, Gleeson M, Fan J;

PI Brasch MA, Cheo D, Hartley JL, Byrd DRN, Temple GF;
XX WPI; 2004-090674/09.
XX Novel isolated nucleic acid molecule comprising recombination sites,
PT topoisomerase recognition sites and/or topoisomerases, useful for in
PT vitro cloning of nucleic acid.
XX Claim 12; Fig 25; 369pp; English.
XX The invention comprises an isolated nucleic acid sequence (e.g. a vector)
CC that contains one or more recombination sites, one or more topoisomerase
CC recognition sites, and/or one or more topoisomerases. The nucleic acid
CC sequence of the invention is useful for in vitro cloning (e.g.
CC recombinatorial cloning) of a nucleic acid molecule. The present DNA
CC sequence represents a plasmid vector of the invention.
XX SQ Sequence 5173 BP; 1247 A; 1337 C; 1274 G; 1305 T; 0 U; 10 Other;
Query Match 100.0%; Score 101; DB 12; Length 5173;
Best Local Similarity 100.0%; Pred. No. 4.3e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CGGCTATTCTTTTGATTATTAAGGGATTTGGGGATTTGGCCCTATTGGTTAAAAAATG 60
Db 1747 CGGCTATTCTTTTGATTATTAAGGGATTTGGGGATTTGGCCCTATTGGTTAAAAAATG 1806
Qy 61 AGCTGATTTAACAAAAATTTAACGCGAATTAATTCGTGGA 101
Db 1807 AGCTGATTTAACAAAAATTTAACGCGAATTAATTCGTGGA 1847
RESULT 10
ID ADM97811 standard; DNA; 5218 BP.
XX ADM97811;
XX 01-JUL-2004 (first entry)
XX & X UAS beta-lactamase vector SEQ ID NO: 64.
XX enzyme; sensor cell; signal transduction detection system; promoter;
KW targeting sequence; targeted drug; ds; vector.
XX Synthetic.
OS Unidentified.
XX WO2004031415-A2.
XX 15-APR-2004.
XX 05-SEP-2003; 2003WO-US028078.
XX 05-SEP-2002; 2002US-0408297P.
XX (VERT-) VERTEX PHARM INC.
XX Whitney MA, Zeh K, Sanders PS;
XX WPI; 2004-330208/30.
XX Developing a sensor cell, useful in determining the activity of a target
PT gene and in developing therapeutic drugs, comprises providing cells
PT comprising a signal transduction detection system and introducing DNA
PT construct into cells.
XX Example 7; Page 231-234; 234pp; English.
XX The present invention relates to a method of developing a sensor cell,
CC for determining the activity of a target gene in the cell, which
CC comprises providing a homogeneous population of cells, where each of the
CC cells comprises a signal transduction detection system and introducing

CC into the population of cells an isolated DNA construct comprising a
CC promoter operatively linked to a targeting sequence. The method is useful
CC in developing a sensor cell for determining the activity of a target gene
CC in the cell. The sensor cell and the methods are useful in developing new
CC and therapeutic drugs directed to the targets. The present sequence is a
CC vector used in the exemplification of the invention.
XX SQ Sequence 5218 BP; 1231 A; 1361 C; 1335 G; 1291 T; 0 U; 0 Other;
Query Match 100.0%; Score 101; DB 12; Length 5218;
Best Local Similarity 100.0%; Pred. No. 4.3e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CGGCTATTCTTTTGATTATTAAGGGATTTGGGGATTTGGCCCTATTGGTTAAAAAATG 60
Db 4624 CGGCTATTCTTTTGATTATTAAGGGATTTGGGGATTTGGCCCTATTGGTTAAAAAATG 4683
Qy 61 AGCTGATTTAACAAAAATTTAACGCGAATTAATTCGTGGA 101
Db 4684 AGCTGATTTAACAAAAATTTAACGCGAATTAATTCGTGGA 4724
RESULT 11
ID ADI34681 standard; DNA; 5302 BP.
XX ADI34681;
XX 22-APR-2004 (first entry)
XX Nucleotide sequence of plasmid pcDNA6/Biotag (TM) -D-TOPO.
XX Recombinational cloning; recombination; topoisomerase;
KW fusion protein production; ds.
XX Synthetic.
XX WO2004005482-A2.
XX 15-JAN-2004.
XX 08-JUL-2003; 2003WO-US021339.
XX 08-JUL-2002; 2002US-0393756P.
PR 19-JUL-2002; 2002US-0396627P.
PR 10-OCT-2002; 2002US-0417172P.
XX (INVI-) INVITROGEN CORP.
XX Bennett RP;
XX WPI; 2004-091356/09.
XX New isolated nucleic acid molecules having one or more recombination
PT sites and encoding an amino acid sequence tag, useful for recombinational
PT and/or topoisomerase-mediated cloning methods for producing fusion
PT proteins.
XX Example 1; Fig 11A-B; 135pp; English.
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC one or more recombination sites, and one or more nucleic acid sequences
CC which encode an amino acid sequence tag. (I) can also comprise one or
CC more topoisomerase recognition sites and/or one or more topoisomerases.
CC The amino acid sequence tag is an amino acid sequence that is capable of
CC being post-translationally modified, and is an amino acid sequence that
CC is capable of being post-translationally modified by biotinylation,
CC attachment of 4-phosphopantetheine, attachment of lipolic acid or
CC attachment of flavins, and is an amino acid sequence that is capable of
CC being biotinylated, wherein the amino acid sequence that is capable of
CC being biotinylated is all or a portion of the Klebsiella pneumoniae
CC oxalacetate decarboxylase a subunit, all or a portion of the
CC Propionibacterium shermanii transcarboxylase 1.3S subunit, or all or a

CC portion of the Escherichia coli biotin carboxyl carrier protein component
 CC of acetyl-CoA carboxylase. The methods and compositions of the present
 CC invention are useful for identifying, concentrating, purifying and/or
 CC producing fusion proteins that comprise an amino acid sequence tag. The
 CC nucleic acid molecules can also be used in recombinational cloning and/or
 CC topoisomerase-mediated cloning methods to produce polynucleotide
 CC constructs which encode the fusion proteins. The present sequence
 CC represents the nucleotide sequence of a plasmid pCDNA6/Biotag(TW)-D-TOPO
 XX Sequence 5302 BP; 1254 A; 1361 C; 1349 G; 1338 T; 0 U; 0 Other;
 SQ
 Query Match 100.0%; Score 101; DB 12; Length 5302;
 Best Local Similarity 100.0%; Pred. No. 4.3e-19;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CGGTCATCTCTTTGATTTATTAAGGATTTTGGGATTTTCGGCCATTGCTTAAATAATG 60
 Db 1872 CGGTCATCTCTTTGATTTATTAAGGATTTTGGGATTTTCGGCCATTGCTTAAATAATG 1931
 QY 61 AGCTGATTTTACAAAAATTAACGCGAATTAATCTGTGGA 101
 Db 1932 AGCTGATTTTACAAAAATTAACGCGAATTAATCTGTGGA 1972

RESULT 12
 ADH11233
 ID ADH11233 standard; DNA; 5425 BP.
 XX
 AC ADH11233;
 XX
 DT 11-MAR-2004 (first entry)
 XX
 DE Vertebrate UNC-53 protein homologue related nucleotide sequence.
 XX
 KW UNC-53 vertebrate protein homologue; UNC-53; Caenorhabditis elegans;
 KW cell shape regulator; cell motility regulator; cell migration;
 KW cell behaviour regulator; phenotype; signal transduction pathway;
 KW signal transducing protein; signal integrator protein;
 KW neuronal regeneration; revascularisation; wound healing;
 KW chronic neurodegenerative disease; acute traumatic injury;
 KW fibrotic disease; gene; ds.
 XX
 OS Unidentified.
 XX
 PN WO9824810-A2.
 XX
 PD 11-JUN-1998.
 XX
 PF 03-DEC-1997; 97WO-EP006956.
 XX
 PR 04-DEC-1996; 96GB-00025283.
 XX
 PA (JANC) JANSSEN PHARM NV.
 XX
 PI Plattreuw CJ, Buesa Arjol CM, Deraeymaeker M, Verhasselt P;
 PI Pujol NJR, Maertens LJS, Luyten W, Geerts H, Vandekerckhove JS;
 PI Geyse J, Bogaert TAOE;
 XX
 DR WPI; 1998-362411/31.
 DR P-PSDB; ADH11234.
 XX
 PT Vertebrate homologues of C. elegans UNC-53 protein - useful for, e.g.
 PT promoting neuronal regeneration, treating chronic neuro-degenerative
 PT diseases or acute traumatic injuries.
 XX
 PS Disclosure; Page 231-237; 479pp; English.
 XX
 CC The present invention describes a vertebrate protein homologue of an UNC-
 CC 53 protein of Caenorhabditis elegans or a functional equivalent,
 CC derivative or bioprecursor of UNC-53. Also described: (1) a cDNA sequence
 CC encoding a vertebrate homologue of the C. elegans UNC-53 protein; (2) a
 CC nucleic acid which hybridises to the cDNA of (1); (3) vector comprising
 CC the cDNA as in (1); (4) a host cell containing the vector as in (3); (5)

CC a transgenic cell, tissue or animal comprising the vector as in (3); (6)
 CC a compound identified as an enhancer or inhibitor of the regulation of
 CC cell shape, motility, or the direction of cell migration for use as a
 CC therapeutic; (7) a method for determination of whether a protein is an
 CC inhibitor or enhancer of regulation of cell behaviour, growth, shape or
 CC motility or the direction of migration by contacting a host cell
 CC expressing a homologue of UNC-53 and determining a change of phenotype;
 CC (8) a method for identification of vertebrate homologues of C. elegans
 CC unc-53 gene by hybridising a probe of 15-50 bp of the unc-53 sequence to
 CC a DNA library; and (9) a method for identification of a protein which is
 CC active in the signal transduction pathway of a cell of which a vertebrate
 CC homologue of UNC-53 is a component comprising: (i) contacting an extract
 CC of a cell with an antibody to the UNC-53 homologue; (ii) identifying an
 CC antibody/homologue complex; and (iii) analysing such a complex to
 CC identify any non-antibody protein bound to the complex. UNC-53 is a
 CC signal transducing or signal integrator protein involved in controlling
 CC directionality of cell migration and cell shape in C. elegans. Vertebrate
 CC homologues of UNC-53 can be used to promote neuronal regeneration,
 CC revascularisation or wound healing, to treat chronic neurodegenerative
 CC diseases or acute traumatic injuries or fibrotic diseases. The present
 CC sequence is used in the exemplification of the present invention.
 XX
 SQ Sequence 5425 BP; 1250 A; 1463 C; 1420 G; 1292 T; 0 U; 0 Other;
 Query Match 100.0%; Score 101; DB 2; Length 5425;
 Best Local Similarity 100.0%; Pred. No. 4.3e-19;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CGGTCATCTCTTTGATTTATTAAGGATTTTGGGATTTTCGGCCATTGCTTAAATAATG 60
 Db 3664 CGGTCATCTCTTTGATTTATTAAGGATTTTGGGATTTTCGGCCATTGCTTAAATAATG 3723
 QY 61 AGCTGATTTTACAAAAATTAACGCGAATTAATCTGTGGA 101
 Db 3724 AGCTGATTTTACAAAAATTAACGCGAATTAATCTGTGGA 3764
 RESULT 13
 ABN86685
 ID ABN86685 standard; DNA; 5431 BP.
 XX
 AC ABN86685;
 XX
 DT 05-NOV-2002 (first entry)
 XX
 DE Nucleotide sequence of a pCDNA3 plasmid vector.
 KW Major histocompatibility complex; MHC; antigen presenting cell; APC;
 KW antigen; cytostatic; virucide; gene therapy; CD8; vaccine; therapeutic;
 KW cancer; viral infection; ds.
 XX
 OS Synthetic.
 XX
 PN WO200261113-A2.
 XX
 PD 08-AUG-2002.
 XX
 PF 01-FEB-2002; 2002WO-US002598.
 XX
 PR 01-FEB-2001; 2001US-0265334P.
 XX
 PA (UJJO) UNIV JOHNS HOPKINS.
 XX
 PI Wu T, Hung C;
 XX
 DR WPI; 2002-619261/66.
 XX
 PT Nucleic acid molecule encoding a fusion polypeptide that promotes
 PT processing via the Major Histocompatibility Complex class I pathway
 PT and/or promotes activity of an antigen presenting cell, useful as vaccine
 PT for cancer and viral infections.
 XX
 PS Claim 24; Page 22-23; 127pp; English.

XX The invention relates to a new nucleic acid molecule (I) encoding a
 CC fusion polypeptide useful as a vaccine composition. (I) comprises a first
 CC nucleic acid sequence encoding a first polypeptide or peptide that
 CC promotes processing via the Major Histocompatibility Complex (MHC) class
 CC I pathway (MHC-I-PP) and/or promotes development or activity of an
 CC antigen presenting cell (APC). The second nucleic acid sequence is linked
 CC in frame to the first nucleic acid sequence or to a linker nucleic acid
 CC sequence and encodes an antigenic polypeptide or peptide. The methods and
 CC compositions of the present invention are useful as therapeutic vaccine
 CC for cancer and for major viral infections, such as hepatoma and cervical
 CC cancer, that cause morbidity and mortality. They can also be used in
 CC treating animal diseases, such as equine herpesvirus, bovine viruses,
 CC Marek's disease, retroviral and lentiviral diseases and rabies, in the
 CC veterinary medicine context. The present sequence represents the
 CC nucleotide sequence of a pcDNA3 plasmid vector
 XX
 SQ Sequence 5431 BP; 1253 A; 1413 C; 1386 G; 1379 T; 0 U; 0 Other;
 Query Match 100.0%; Score 101; DB 6; Length 5431;
 Best Local Similarity 100.0%; Pred. No. 4.3e-19;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CGGTCATTCTTTGATTATTAAGGGATTTGGGGATTTGGCCCTATTGGTTAAAAAATG 60
 |||||
 Db 1636 CGGTCATTCTTTGATTATTAAGGGATTTGGGGATTTGGCCCTATTGGTTAAAAAATG 1695
 QY 61 AGCTGATTAAACAAAAATTTAACGCGAATTAAATCTGTGGA 101
 |||||
 Db 1696 AGCTGATTAAACAAAAATTTAACGCGAATTAAATCTGTGGA 1736
 RESULT 14
 ADE21866
 ID ADE21866 standard; DNA; 5431 BP.
 AC ADE21866;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Plasmid vector pcDNA3 nucleotide sequence SEQ ID NO:8.
 XX
 KW chimeric fusion; translocation; antigenic; cytostatic; immunotherapy;
 KW gene therapy; cancer; tumour; gene; ds.
 XX
 OS Synthetic.
 XX
 PN WO2003085085-A2.
 XX
 PD 16-OCT-2003.
 XX
 PF 04-APR-2003; 2003WO-US010235.
 XX
 PR 04-APR-2002; 2002US-00115440.
 XX
 PA (UIJO) UNIV JOHNS HOPKINS.
 XX
 PI Wu T, Hung C;
 XX
 PI WPI; 2003-877027/81.
 DR
 XX
 PT New nucleic acid encoding a chimeric fusion or fusion polypeptide
 PT comprising a first domain with a translocation polypeptide, and a second
 PT domain with an antigen having at least one antigenic peptide, useful for
 PT treating cancer.
 PT
 PS Disclosure; SEQ ID NO 8; 68pp; English.
 XX
 CC The present invention describes a nucleic acid (I) encoding a chimeric
 CC fusion or fusion polypeptide comprising a first domain with a
 CC translocation polypeptide, and a second domain comprising an antigen
 CC having at least one antigenic peptide. Also described: (i) an expression
 CC vector comprising (I) operatively linked to a promoter and optionally, to

CC one or more regulatory elements that enhance expression of the nucleic
 CC acid in a cell; (2) a particle comprising (I) or the expression vector;
 CC (3) a cell that has been modified to comprise (I) or the expression
 CC vector; (4) a chimeric polypeptide comprising a first domain with a
 CC translocation polypeptide, and a second domain comprising an antigen
 CC having at least one antigenic peptide; (5) a pharmaceutical composition
 CC capable of inducing or enhancing an antigen specific immune response,
 CC comprising (I), expression vector, particle, cell, cell of the particle,
 CC or the chimeric polypeptide; and a carrier or excipient; (6) inducing or
 CC enhancing an antigen specific immune response by administering the
 CC composition described above; (7) increasing the number of CD8 + CTLs
 CC specific for a selected desired antigen in a subject by administering the
 CC composition described above; and (8) inhibiting the growth of a tumour in
 CC a subject by administering the composition described above. (I) has
 CC cytostatic activity, and can be used in immunotherapy, and gene therapy.
 CC The nucleic acids (I), compositions and methods are useful for treating
 CC cancer. The present sequence represents a plasmid vector nucleotide
 CC sequence which is used in the exemplification of the present invention.
 XX
 SQ Sequence 5431 BP; 1253 A; 1413 C; 1386 G; 1379 T; 0 U; 0 Other;
 Query Match 100.0%; Score 101; DB 10; Length 5431;
 Best Local Similarity 100.0%; Pred. No. 4.3e-19;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CGGTCATTCTTTGATTATTAAGGGATTTGGGGATTTGGCCCTATTGGTTAAAAAATG 60
 |||||
 Db 1636 CGGTCATTCTTTGATTATTAAGGGATTTGGGGATTTGGCCCTATTGGTTAAAAAATG 1695
 QY 61 AGCTGATTAAACAAAAATTTAACGCGAATTAAATCTGTGGA 101
 |||||
 Db 1696 AGCTGATTAAACAAAAATTTAACGCGAATTAAATCTGTGGA 1736
 RESULT 15
 ADO05277
 ID ADO05277 standard; DNA; 5431 BP.
 XX
 AC ADO05277;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE pcDNA3 plasmid vector.
 XX
 KW Translocation domain; bacterial toxin; exotoxin A domain II; ETA;
 KW major histocompatibility complex; MHC class I; vaccine; immune response;
 KW CD8+ cytotoxic T lymphocyte; CTL; tumour; E7 antigen; pcDNA3 plasmid; ds.
 XX
 OS Synthetic.
 XX
 PN US2004086845-A1.
 XX
 PD 06-MAY-2004.
 XX
 PF 04-APR-2002; 2002US-00115440.
 XX
 PR 20-OCT-1999; 99US-00421608.
 PR 09-FEB-2000; 2000US-00501097.
 PR 20-OCT-2000; 2000WO-US041422.
 PR 04-APR-2001; 2001US-0281003P.
 XX
 PA (WUTY/) WU T.
 PA (HUNG/) HUNG C.
 XX
 PI Wu T, Hung C;
 XX
 PI WPI; 2004-356187/33.
 DR
 XX
 PT Novel chimeric polypeptide e.g., Pseudomonas aeruginosa exotoxin A domain
 PT II/human papilloma virus-16 E7 peptide useful for inducing or enhancing
 PT antigen specific immune response, or for inhibiting growth of tumor in
 PT subject.

PS Disclosure; SEQ ID NO 8; 48pp; English.

XX
CC The invention relates to nucleic acid encoding a chimeric polypeptide
CC comprising a translocation domain of a bacterial toxin and at least one
CC antigenic peptide. The preferred translocation domain is domain II of
CC pseudomonas aeruginosa exotoxin A (ETA(dII)) and the preferred antigen is
CC human papilloma virus type 16 (HPV-16) E7 which is a model tumour
CC antigen. The antigenic peptide comprises an epitope that binds to and is
CC presented on the cell surface by major histocompatibility complex (MHC)
CC class I proteins. The nucleic acid of the invention is useful as vaccine
CC composition for enhancing antigen specific immune response, increasing
CC the number of CD8+ cytotoxic T lymphocytes (CTLs) and for inhibiting the
CC growth of a tumour. The present sequence is pcDNA3 plasmid vector used in
CC the invention.

XX
SQ Sequence 5431 BP; 1253 A; 1413 C; 1386 G; 1379 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 12; Length 5431;

Best Local Similarity 100.0%; Pred. No. 4.3e-19;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGTCATTCTTTGATTTATAGGGATTTTCGGGATTTTCGGCTATTGTTAAAAAATG 60

Db 1636 CGGTCATTCTTTGATTTATAGGGATTTTCGGGATTTTCGGCTATTGTTAAAAAATG 1695

QY 61 AGCTGATTTAACAAAAATTTAACGGGAATTAATTCGTGGA 101

Db 1696 AGCTGATTTAACAAAAATTTAACGGGAATTAATTCGTGGA 1736

Search completed: July 14, 2005, 07:01:25

Job time : 143.448 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 05:15:57 ; Search time 961.667 Seconds
(without alignments)
3997.736 Million cell updates/sec

Title: US-09-482-682-8_COPY_3674_3774
Perfect score: 101
Sequence: 1 cggctctattctttgattta.....acgcgaattaattctgtgga 101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	88.4	87.5	827	5	BQ151093 NF044H03L
C 2	88.2	87.3	805	5	BQ145762 NF017H10G
C 3	88.2	87.3	813	5	BQ145733 NF017D06G
C 4	88	87.1	585	5	BQ998405 HI10P06r
C 5	87.4	86.5	750	5	BQ144787 NF104G07D
C 6	87.4	86.5	750	5	BQ154917 NF074C04i
C 7	87.4	86.5	764	5	BQ155881 NF085C08i
C 8	87.4	86.5	768	5	BQ159395 NF17H07P
C 9	87.4	86.5	783	5	BQ158754 NF070H02P
C 10	87.4	86.5	784	5	BQ159157 NF05D02P
C 11	87.4	86.5	784	5	BQ159352 NF114D02P
C 12	87.4	86.5	785	5	BQ156587 NF094D10i
C 13	87.4	86.5	796	5	BQ139350 NF014C09P
C 14	87.4	86.5	796	5	BQ153647 NF040H05i
C 15	87.4	86.5	798	5	BQ154182 NF056D06i
C 16	87.4	86.5	816	5	BQ144358 NF072C01D
C 17	87.4	86.5	820	5	BQ154225 NF056C02i
C 18	87.4	86.5	832	5	BQ146006 NF024A04G
C 19	87.2	86.3	389	4	B1938184 de35h05.y
C 20	87.2	86.3	392	4	B1938175 de35g04.y
C 21	87.2	86.3	395	6	CA620718 wlln.pk00
C 22	87.2	86.3	453	6	CA624892 wlln.pk01
C 23	87.2	86.3	472	4	BM276444 PfEST0a8
C 24	87.2	86.3	565	6	CA623376 wlln.pk01

25	87.2	86.3	588	6	CA617091
26	86.8	85.9	137	6	CD282765
27	86.8	85.9	147	6	CD285004
C 28	86.8	85.9	154	6	CD288994
29	86.8	85.9	161	7	CN523248
30	86.8	85.9	164	6	CD282911
C 31	86.8	85.9	181	8	AQ014620
C 32	86.8	85.9	183	8	AQ014434
C 33	86.8	85.9	185	2	BF942500
C 34	86.8	85.9	188	8	B97287
35	86.8	85.9	193	6	CD281345
C 36	86.8	85.9	193	8	BZ666758
37	86.8	85.9	202	6	CD280039
C 38	86.8	85.9	204	8	B68786
C 39	86.8	85.9	205	6	CD280392
40	86.8	85.9	207	7	CN523315
41	86.8	85.9	208	7	CN521451
C 42	86.8	85.9	210	1	AJ409361
C 43	86.8	85.9	217	8	B60716
C 44	86.8	85.9	224	6	CD280393
45	86.8	85.9	235	6	CD280049

ALIGNMENTS

RESULT 1
BQ151093/C
LOCUS BQ151093.1 827 bp mRNA linear EST 24-APR-2002
DEFINITION NF044H03L.F1029 Developing leaf Medicago truncatula cDNA clone
ACCESSION BQ151093
VERSION BQ151093.1 GI:20288152
KEYWORDS EST.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula

REFERENCE 1 (bases 1 to 827)
AUTHORS Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J., Flores,H.R., Inman,J.T., Weller,J.W. and May,G.D.
TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation Medicago truncatula leaf library
JOURNAL Unpublished (2000)
COMMENT Contact: May GD
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 224 6650
Fax: 580 224 6692
Email: gdmay@noble.org
Insert Length: 827 Std Error: 0.00
Plate: 044 row: H column: 03
Seq primer: TCACACGAGAAACAGCTATGAC.
Location/Qualifiers
1. .827
/organism="Medicago truncatula"
/mol_type="mRNA"
/db_xref="taxon:3880"
/clone="NF044H03L.F"
/tissue_type="leaf"
/dev_stage="Pooled developmental"
/clone_lib="Developing leaf"
/note="Vector: Lambda Zap; Contains a mixture of very young, developing, mature and senescing leaves."

Query Match 87.5%; Score 88.4; DB 5; Length 827;
Best Local Similarity 98.9%; Pred. No. 1.1e-13;
Matches 89; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGTCTATCTTTGATTATATAAGGATTTTGGGATTTTCGCCCTATTGGTTAAAAAATG 60
 Db |||||||
 483 CGGTCTATCTTTGATTATATAAGGATTTTGGGATTTTCGCCCTATTGGTTAAAAAATG 424
 |||||||

QY 61 AGCTGATTAAACAAAAATTTAACGCGAATT 90
 Db |||||||

423 AGCTGATTAAACAAAAATTTAACGCGAATT 394
 |||||||

RESULT 2
 BQ145762/c
 LOCUS
 DEFINITION NF017H10GS1F1091 Germinating Seed Medicago truncatula cDNA clone
 NF017H10GS 5', mRNA sequence.
 ACCESSION BQ145762
 VERSION BQ145762.1 GI:20282821
 KEYWORDS EST.
 SOURCE Medicago truncatula (barrel medic)
 ORGANISM Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
 Medicago.

REFERENCE
 1 (bases 1 to 805)
 Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,
 Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.
 Expressed Sequence Tags from the Samuel Roberts Noble Foundation
 Medicago truncatula germinating seed library
 Unpublished (2001)

JOURNAL
 COMMENT Plant Biology Division
 The Samuel Roberts Noble Foundation
 2510 Sam Noble Parkway, Ardmore, OK 73402, USA
 Tel: 580 224 6650
 Fax: 580 224 6692
 Email: gdmay@noble.org
 Insert Length: 805 Std Error: 0.00
 Plate: 017 row: H column: 10
 Seq primer: TCACACAGGAACAGCTATGAC.
 Location/Qualifiers

FEATURES
 source
 1..805

/organism="Medicago truncatula"
 /mol_type="mRNA"
 /db_xref="taxon:3880"
 /clone="NF017H10GS"
 /tissue_type="germinating seeds"
 /dev_stage="0, 1, 2 and 3 days after acid treatment."
 /clone_lib="Germinating Seed"
 /note="Vector: Lambda Zap; M. truncatula seeds were acid
 treated and placed on wet filter papers in petri dishes.
 Seeds were harvested at 0, 1, 2 and 3 days after acid
 treatment. cDNA was prepared from polyA+ enriched, pooled
 samples of equivalent amounts of total RNA from each time
 point. The cDNA was directionally ligated into the
 Uni-Zap XR vector (Stratagene) and packaged using the
 Gigapack III Gold packaging extracts. Phagemids
 containing cDNA inserts were in vivo excised from the
 recombinant Uni-Zap XR vector using ExAssist helper phage
 and the E. coli strain XLI-Blue MRF' (Stratagene).
 Excised plasmids were plated using SOLR cells."

ORIGIN
 Query Match 87.3%; Score 88.2; DB 5; Length 805;
 Best Local Similarity 92.1%; Pred. NO. 1.3e-13;
 Matches 93; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CGGTCTATCTTTGATTATATAAGGATTTTGGGATTTTCGCCCTATTGGTTAAAAAATG 60
 Db 232 CGGTCTATCTTTGATTATATAAGGATTTTGGGATTTTCGCCCTATTGGTTAAAAAATG 173
 |||||||

QY 61 AGCTGATTAAACAAAAATTTAACGCGAATTCTCTGGA 101
 Db 172 AGCTGATTAAACAAAAATTTAACGCGAATTCCTCGTCCGA 132
 |||||||

RESULT 3
 BQ145733/c

LOCUS
 DEFINITION NF017D06GS1F1057 Germinating Seed Medicago truncatula cDNA clone
 NF017D06GS 5', mRNA sequence.

ACCESSION BQ145733
 VERSION BQ145733.1 GI:20282792
 KEYWORDS EST.
 SOURCE Medicago truncatula (barrel medic)
 ORGANISM Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
 Medicago.

REFERENCE
 1 (bases 1 to 813)
 Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,
 Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.
 Expressed Sequence Tags from the Samuel Roberts Noble Foundation
 Medicago truncatula germinating seed library
 Unpublished (2001)

JOURNAL
 COMMENT Plant Biology Division
 The Samuel Roberts Noble Foundation
 2510 Sam Noble Parkway, Ardmore, OK 73402, USA
 Tel: 580 224 6650
 Fax: 580 224 6692
 Email: gdmay@noble.org
 Insert Length: 813 Std Error: 0.00
 Plate: 017 row: D column: 06
 Seq primer: TCACACAGGAACAGCTATGAC.
 Location/Qualifiers

FEATURES
 source
 1..813

/organism="Medicago truncatula"
 /mol_type="mRNA"
 /db_xref="taxon:3880"
 /clone="NF017D06GS"
 /tissue_type="germinating seeds"
 /dev_stage="0, 1, 2 and 3 days after acid treatment."
 /clone_lib="Germinating Seed"
 /note="Vector: Lambda Zap; M. truncatula seeds were acid
 treated and placed on wet filter papers in petri dishes.
 Seeds were harvested at 0, 1, 2 and 3 days after acid
 treatment. cDNA was prepared from polyA+ enriched, pooled
 samples of equivalent amounts of total RNA from each time
 point. The cDNA was directionally ligated into the
 Uni-Zap XR vector (Stratagene) and packaged using the
 Gigapack III Gold packaging extracts. Phagemids
 containing cDNA inserts were in vivo excised from the
 recombinant Uni-Zap XR vector using ExAssist helper phage
 and the E. coli strain XLI-Blue MRF' (Stratagene).
 Excised plasmids were plated using SOLR cells."

ORIGIN

Query Match 87.3%; Score 88.2; DB 5; Length 813;
 Best Local Similarity 92.1%; Pred. No. 1.3e-13;
 Matches 93; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CGGTCTATCTTTGATTATATAAGGATTTTGGGATTTTCGCCCTATTGGTTAAAAAATG 60
 Db 231 CGGTCTATCTTTGATTATATAAGGATTTTGGGATTTTCGCCCTATTGGTTAAAAAATG 172
 |||||||

QY 61 AGCTGATTAAACAAAAATTTAACGCGAATTAAATCTCTGGA 101
 Db 171 AGCTGATTAAACAAAAATTTAACGCGAATTCCTCGTCCGA 131
 |||||||

RESULT 4
 BU998405
 LOCUS
 DEFINITION HI10P06r HI Hordeum vulgare subsp. vulgare cDNA clone HI10P06
 5-PRIME, mRNA sequence.

BU998405 585 bp mRNA linear EST 23-OCT-2002
 HI10P06r HI Hordeum vulgare subsp. vulgare cDNA clone HI10P06
 5-PRIME, mRNA sequence.

```

ACCESSION   BU998405
VERSION      BU998405.1  GI:24275388
KEYWORDS     EST.
SOURCE       ORGANISM
             Hordeum vulgare subsp. vulgare
             Hordeum vulgare subsp. vulgare
             Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
             Pooidae; Triticeae; Hordeum.
REFERENCE    1 (bases 1 to 585)
AUTHORS      Zhang,H., Weschke,W., Michalek,W., Stein,N. and Graner,A.
TITLE        EST sequencing and analysis in barley (2002)
JOURNAL      Unpublished (2002)
COMMENT      Contact: Stein Nils
             Molecular Markers Group, Department Genbank
             Institute of Plant Genetics and Crop Plant Research (IPK)
             Corrensstr. 3, 06466, Gatersleben, Germany
             Tel: 039482-5522
             Fax: 039482-5595
             Email: stein@ipk-gatersleben.de
             Insert Length: 585 Std Error: 0.00
             Plate: 10 row: P column: 6
             Seq primer: M13rev.
FEATURES     source
             Location/Qualifiers
             1..585
             /organism="Hordeum vulgare subsp. vulgare"
             /mol_type="mRNA"
             /cultivar="barke"
             /sub_species="vulgare"
             /db_xref="GABI:252269"
             /db_xref="taxon:112509"
             /clone="H110P06"
             /tissue_type="female inflorescences"
             /dev_stage="female inflorescences (approx. 3 mm in size)"
             /lab_host="XL10-Gold"
             /clone_lib="H1"
             /note="Vector: pBluescript SK+; Site 1: EcoRI (5'-end of
             cDNA); Site 2: XbaI (3'-end of cDNA); Due to a cloning
             artefact caused by the kit, in most cases the EcoRI site
             is NOT present, as well as the EcoRI adapter used for
             cloning. To excise the insert, restriction sites upstream
             EcoRI should be used (e.g. BamHI, SalI, PstI). NOTE: Also
             due to the cloning system used Blue/white selection for
             recombinants is not 100% reliable."
ORIGIN
Query Match      87.1%; Score 88; DB 5; Length 585;
Best Local Similarity 94.8%; Pred. No. 1.4e-13;
Matches 91; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGGCTATTCTTTGATTATTAAGGGATTTGGGGATTTCCGCCCTATTGGTTAAAAAATG 60
    |||||
Db 54 CGGCTATTCTTTGATTATTAAGGGATTTGGGGATTTCCGCCCTATTGGTTAAAAAATG 113
    |||||

Qy 61 AGCTGATTAAACAAAAATTTAACCGGAATTAATCT 96
    |||||
Db 114 AGCTGATTAAACAAAAATTTAACCGGAATTTT 149
    |||||

RESULT 5
BQ144787/c
LOCUS       BQ144787 750 bp mRNA linear EST 24-APR-2002
DEFINITION NF104G07D7LF1054 Drought Medicago truncatula cDNA clone NF104G07D
5', mRNA sequence.
ACCESSION   BQ144787
VERSION     BQ144787.1  GI:20281846
KEYWORDS    EST.
SOURCE      Medicago truncatula (barrel medic)
ORGANISM    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
             rosids; euroids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
             Medicago.
             1 (bases 1 to 750)
REFERENCE    1 (bases 1 to 750)
AUTHORS      Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
             Flores,H.R., Imman,J.T., Weller,J.W. and May,G.D.
TITLE        Expressed Sequence Tags from the Samuel Roberts Noble Foundation
JOURNAL      Unpublished (2001)
COMMENT      Contact: May GD
             Plant Biology Division
             The Samuel Roberts Noble Foundation
             2510 Sam Noble Parkway, Ardmore, OK 73402, USA
             Tel: 580 224 6650
             Fax: 580 224 6692
             Email: gdmay@noble.org
             Insert Length: 750 Std Error: 0.00
             Plate: 074 row: C column: 04
             Seq primer: TCACACAGGAACAGCTATGAC.
FEATURES     source
             Location/Qualifiers
             1..750
             /organism="Medicago truncatula"
             /mol_type="mRNA"
             /db_xref="taxon:3880"
             /clone="NF104G07DT"
             /tissue_type="Plantlets"
             /dev_stage="Pooled timepoints"
             /clone_lib="Drought"
             /note="vector: Lambda Zap; Contains a mixture of entire
             plantlets harvested in a series of days-post-watering
             timepoints."
ORIGIN
Query Match      86.5%; Score 87.4; DB 5; Length 750;
Best Local Similarity 93.8%; Pred. No. 2.1e-13;
Matches 91; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CGGCTATTCTTTGATTATTAAGGGATTTGGGGATTTCCGCCCTATTGGTTAAAAAATG 60
    |||||
Db 192 CGGCTATTCTTTGATTATTAAGGGATTTGGGGATTTCCGCCCTATTGGTTAAAAAATG 133
    |||||

Qy 61 AGCTGATTAAACAAAAATTTAACCGGAATTAATCTG 97
    |||||
Db 132 AGCTGATTAAACAAAAATTTAACCGGAATTCCTGAG 96
    |||||

RESULT 6
BQ154917/c
LOCUS       BQ154917 750 bp mRNA linear EST 24-APR-2002
DEFINITION NF074C041R1F1034 Irradiated Medicago truncatula cDNA clone
NF074C041R 5', mRNA sequence.
ACCESSION   BQ154917
VERSION     BQ154917.1  GI:20291976
KEYWORDS    EST.
SOURCE      Medicago truncatula (barrel medic)
ORGANISM    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
             rosids; euroids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
             Medicago.
             1 (bases 1 to 750)
REFERENCE    1 (bases 1 to 750)
AUTHORS      Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
             Flores,H.R., Imman,J.T., Weller,J.W. and May,G.D.
TITLE        Expressed Sequence Tags from the Samuel Roberts Noble Foundation
JOURNAL      Unpublished (2001)
COMMENT      Contact: May GD
             Plant Biology Division
             The Samuel Roberts Noble Foundation
             2510 Sam Noble Parkway, Ardmore, OK 73402, USA
             Tel: 580 224 6650
             Fax: 580 224 6692
             Email: gdmay@noble.org
             Insert Length: 750 Std Error: 0.00
             Plate: 074 row: C column: 04
             Seq primer: TCACACAGGAACAGCTATGAC.
FEATURES     source
             Location/Qualifiers
             1..750
             /organism="Medicago truncatula"
             /mol_type="mRNA"
             /db_xref="taxon:3880"
             /clone="NF104G07DT"
             /tissue_type="Plantlets"
             /dev_stage="Pooled timepoints"
             /clone_lib="Drought"
             /note="vector: Lambda Zap; Contains a mixture of entire
             plantlets harvested in a series of days-post-watering
             timepoints."
ORIGIN
Query Match      86.5%; Score 87.4; DB 5; Length 750;
Best Local Similarity 93.8%; Pred. No. 2.1e-13;
Matches 91; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CGGCTATTCTTTGATTATTAAGGGATTTGGGGATTTCCGCCCTATTGGTTAAAAAATG 60
    |||||
Db 192 CGGCTATTCTTTGATTATTAAGGGATTTGGGGATTTCCGCCCTATTGGTTAAAAAATG 133
    |||||

Qy 61 AGCTGATTAAACAAAAATTTAACCGGAATTAATCTG 97
    |||||
Db 132 AGCTGATTAAACAAAAATTTAACCGGAATTCCTGAG 96
    |||||

```

```

source
1. .750
/organism="Medicago truncatula"
/mol_type="mRNA"
/db_xref="taxon:3880"
/clone="NF074C041R"
/tissue_type="seedlings"
/dev_stage="seedling"
/clone_lib="Irradiated"
/notes="Vector: Lambda Zap; Seedlings were exposed either
to 100 Gy gamma or 0.5, 1, 5, or 10 kJ/m2 UV irradiation.
Gamma-irradiated samples were harvested at 6, 12, 24 and
48 hours after treatment. UV-irradiated samples were
harvested 24 hours post-treatment. cDNA was prepared from
polyA+ enriched, pooled samples of equivalent amounts of
total RNA from each sample. The cDNA was directionally
ligated into the Uni-Zap XR vector (Stratagene) and
packaged using the Gigapack III Gold packaging extracts.
Phagemids containing cDNA inserts were in vivo excised
from the recombinant Uni-Zap XR vector using ExAassist
helper phage and the E. coli strain XL1-Blue MRF'
(Stratagene). Excised plasmids were plated using SOLR
cells."

ORIGIN
Query Match      86.5%; Score 87.4; DB 5; Length 750;
Best Local Similarity 93.8%; Pred. No. 2.1e-13;
Matches 91; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CGGTCATCTCTTTGATTATATAAGGGATTTCGGGCTATTGGCTATTGGTTAAAAAATG 60
    |||||||
Db 191 CGGTCATCTCTTTGATTATATAAGGGATTTCGGGCTATTGGCTATTGGTTAAAAAATG 132

Qy 61 AGCTGATTTAACAAAAATTTAACGCGAATTAAATCTG 97
    |||||||
Db 131 AGCTGATTTAACAAAAATTTAACGCGAATTCTCTGCAG 95

RESULT 7
BQ155881/c
LOCUS
DEFINITION
  NF085C081R1F1066 Irradiated Medicago truncatula cDNA clone
ACCESSION
  BQ155881
VERSION
  BQ155881.1 GI:20292940
KEYWORDS
  EST.
SOURCE
  Medicago truncatula (barrel medic)
ORGANISM
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;
  rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
  Medicago.
REFERENCE
  1 (bases 1 to 764)
  Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
  Flores,H.R., Inman,J.T., Weller,J.W. and May,G.D.
  Expressed Sequence Tags from the Samuel Roberts Noble Foundation
  Medicago truncatula irradiated library
  Unpublished (2001)
  Contact: May GD
  Plant Biology Division
  The Samuel Roberts Noble Foundation
  2510 Sam Noble Parkway, Ardmore, OK 73402, USA
  Tel: 580 224 6650
  Fax: 580 224 6692
  Email: gdmay@noble.org
  Insert Length: 764 Std Error: 0.00
  Plate: 085 row: C column: 08
  Seq primer: TCACACGGAACAGCTATGAC.
  Location/Qualifiers
    1. .764
    /organism="Medicago truncatula"
    /mol_type="mRNA"
    /db_xref="taxon:3880"
    /clone="NF085C081R"

source
1. .750
/organism="Medicago truncatula"
/mol_type="mRNA"
/db_xref="taxon:3880"
/clone="NF074C041R"
/tissue_type="seedlings"
/dev_stage="seedling"
/clone_lib="Irradiated"
/notes="Vector: Lambda Zap; Seedlings were exposed either
to 100 Gy gamma or 0.5, 1, 5, or 10 kJ/m2 UV irradiation.
Gamma-irradiated samples were harvested at 6, 12, 24 and
48 hours after treatment. UV-irradiated samples were
harvested 24 hours post-treatment. cDNA was prepared from
polyA+ enriched, pooled samples of equivalent amounts of
total RNA from each sample. The cDNA was directionally
ligated into the Uni-Zap XR vector (Stratagene) and
packaged using the Gigapack III Gold packaging extracts.
Phagemids containing cDNA inserts were in vivo excised
from the recombinant Uni-Zap XR vector using ExAassist
helper phage and the E. coli strain XL1-Blue MRF'
(Stratagene). Excised plasmids were plated using SOLR
cells."

ORIGIN
Query Match      86.5%; Score 87.4; DB 5; Length 750;
Best Local Similarity 93.8%; Pred. No. 2.1e-13;
Matches 91; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CGGTCATCTCTTTGATTATATAAGGGATTTCGGGCTATTGGCTATTGGTTAAAAAATG 60
    |||||||
Db 191 CGGTCATCTCTTTGATTATATAAGGGATTTCGGGCTATTGGCTATTGGTTAAAAAATG 132

Qy 61 AGCTGATTTAACAAAAATTTAACGCGAATTAAATCTG 97
    |||||||
Db 131 AGCTGATTTAACAAAAATTTAACGCGAATTCTCTGCAG 95

RESULT 8
BQ159395/c
LOCUS
DEFINITION
  NF117H07PL1F1064 Phosphate starved leaf Medicago truncatula cDNA
ACCESSION
  BQ159395
VERSION
  BQ159395.1 GI:20296452
KEYWORDS
  EST.
SOURCE
  Medicago truncatula (barrel medic)
ORGANISM
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
  Medicago.
REFERENCE
  1 (bases 1 to 768)
  Liu,J., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
  Flores,H.R., Inman,J.T., Weller,J.W., May,G.D. and Harrison,M.J.
  Expressed Sequence Tags from the Samuel Roberts Noble Foundation
  Medicago truncatula phosphate-starved leaf library
  Unpublished (2000)
  Contact: Harrison MJ
  Plant Biology Division
  The Samuel Roberts Noble Foundation
  2510 Sam Noble Parkway, Ardmore, OK 73402, USA
  Tel: 580 221 7325
  Fax: 580 221 7380
  Email: mjharrison@noble.org
  Insert Length: 768 Std Error: 0.00
  Plate: 117 row: H column: 07
  Seq primer: TCACACGGAACAGCTATGAC.
  Location/Qualifiers
    1. .768
    /organism="Medicago truncatula"
    /mol_type="mRNA"
    /db_xref="taxon:3880"
    /clone="NF117H07PL"
    /tissue_type="leaf"
    /dev_stage="trifoliolate"
    /clone_lib="Phosphate starved leaf"
    /note="Vector: Lambda Zap; At the trifoliolate stage, M.
    truncatula plants were transplanted to phosphate-free sand

```

and grown for a further 30 days. During this 30 day period, the plants were fertilized twice weekly with 1/2 Hoaglands solution containing only 20uM potassium phosphate. RNA was prepared from above ground tissues."

ORIGIN

Query Match 86.5%; Score 87.4; DB 5; Length 768;
Best Local Similarity 93.8%; Pred. No. 2.1e-13;
Matches 91; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CGGCTATTCTTTGATTATTAAGGATTTGGGGATTTGGCCCTATTGGTTAAAAAATG 60
Db 193 CGGCTATTCTTTGATTATTAAGGATTTGGCCGATTTGGCCCTATTGGTTAAAAAATG 134

Qy 61 AGCTGATTACAAAAATTTAACGGAATTAATCTG 97
Db 133 AGCTGATTACAAAAATTTAACGGAATTCCTGCAG 97

RESULT 9

BQ158754/c
LOCUS BQ158754 783 bp mRNA linear EST 24-APR-2002
DEFINITION NF070H02PL1F1026 Phosphate starved leaf Medicago truncatula cDNA
clone NF070H02PL 5', mRNA sequence.

ACCESSION BQ158754
VERSION BQ158754.1 GI:20295811
KEYWORDS EST.

SOURCE Medicago truncatula (barrel medic)

ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.

REFERENCE 1 (bases 1 to 783)

AUTHORS Liu, J., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,
Flores, H.R., Imman, J.T., Weller, J.W., May, G.D. and Harrison, M.J.
TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation
Medicago truncatula phosphate-starved leaf library

JOURNAL

COMMENT Unpublished (2000)

CONTACT: Harrison MJ
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7325
Fax: 580 221 7380

Email: mjharrison@noble.org
Insert Length: 783 Std Error: 0.00

Plate: 070 row: H column: 02
Seq primer: TCACACGGAACAGCTATGAC.

FEATURES

Location/Qualifiers
1..783
/organism="Medicago truncatula"
/mol_type="mRNA"
/db_xref="taxon:3880"
/clone="NF070H02PL"
/tissue_type="leaf"
/dev_stage="trifoliolate"

/clone_lib="Phosphate starved leaf"
/note="Vector: Lambda Zap; At the trifoliolate stage, M.
truncatula plants were transplanted to phosphate-free sand
and grown for a further 30 days. During this 30 day
period, the plants were fertilized twice weekly with 1/2
Hoaglands solution containing only 20uM potassium
phosphate. RNA was prepared from above ground tissues."

ORIGIN

Query Match 86.5%; Score 87.4; DB 5; Length 783;
Best Local Similarity 93.8%; Pred. No. 2.1e-13;
Matches 91; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CGGCTATTCTTTGATTATTAAGGATTTGGGGATTTGGCCCTATTGGTTAAAAAATG 60
Db 192 CGGCTATTCTTTGATTATTAAGGATTTGGCCGATTTGGCCCTATTGGTTAAAAAATG 133

Qy 61 AGCTGATTACAAAAATTTAACGGAATTAATCTG 97
Db 132 AGCTGATTACAAAAATTTAACGGAATTCCTGCAG 96

RESULT 10
BQ159157/c

LOCUS BQ159157 784 bp mRNA linear EST 24-APR-2002
DEFINITION NF050D02PL1F1014 Phosphate starved leaf Medicago truncatula cDNA
clone NF050D02PL 5', mRNA sequence.

ACCESSION BQ159157
VERSION BQ159157.1 GI:20296214

KEYWORDS EST.

SOURCE Medicago truncatula (barrel medic)

ORGANISM Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.

REFERENCE 1 (bases 1 to 784)

AUTHORS Liu, J., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J.,
Flores, H.R., Imman, J.T., Weller, J.W., May, G.D. and Harrison, M.J.
TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation
Medicago truncatula phosphate-starved leaf library

JOURNAL

COMMENT Unpublished (2000)

CONTACT: Harrison MJ

Plant Biology Division

The Samuel Roberts Noble Foundation

2510 Sam Noble Parkway, Ardmore, OK 73402, USA

Tel: 580 221 7325

Fax: 580 221 7380

Email: mjharrison@noble.org

Insert Length: 784 Std Error: 0.00

Plate: 050 row: D column: 02

Seq primer: TCACACGGAACAGCTATGAC.

FEATURES

Location/Qualifiers
1..784

/organism="Medicago truncatula"

/mol_type="mRNA"

/db_xref="taxon:3880"

/clone="NF050D02PL"

/tissue_type="leaf"

/dev_stage="trifoliolate"

/clone_lib="Phosphate starved leaf"

/note="Vector: Lambda Zap; At the trifoliolate stage, M.
truncatula plants were transplanted to phosphate-free sand
and grown for a further 30 days. During this 30 day
period, the plants were fertilized twice weekly with 1/2
Hoaglands solution containing only 20uM potassium
phosphate. RNA was prepared from above ground tissues."

ORIGIN

Query Match 86.5%; Score 87.4; DB 5; Length 784;
Best Local Similarity 93.8%; Pred. No. 2.1e-13;
Matches 91; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CGGCTATTCTTTGATTATTAAGGATTTGGGGATTTGGCCCTATTGGTTAAAAAATG 60
Db 192 CGGCTATTCTTTGATTATTAAGGATTTGGCCGATTTGGCCCTATTGGTTAAAAAATG 133

Qy 61 AGCTGATTACAAAAATTTAACGGAATTAATCTG 97
Db 132 AGCTGATTACAAAAATTTAACGGAATTCCTGCAG 96

RESULT 11

BQ159352/c
LOCUS BQ159352 784 bp mRNA linear EST 24-APR-2002
DEFINITION NF114D02PL1F1026 Phosphate starved leaf Medicago truncatula cDNA
clone NF114D02PL 5', mRNA sequence.

ACCESSION BQ159352

VERSION BQ159352.1 GI:20296409

```

KEYWORDS      EST.
SOURCE         Medicago truncatula (barrel medic)
ORGANISM       Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
                Medicago.
REFERENCE      1 (bases 1 to 784)
AUTHORS        Liu,J., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
                Flores,H.R., Inman,J.T., Weller,J.W., May,G.D. and Harrison,M.J.
TITLE          Expressed Sequence Tags from the Samuel Roberts Noble Foundation
                Medicago truncatula phosphate-starved leaf library
JOURNAL        Unpublished (2000)
COMMENT        Contact: Harrison MJ
                Plant Biology Division
                The Samuel Roberts Noble Foundation
                2510 Sam Noble Parkway, Ardmore, OK 73402, USA
                Tel: 580 221 7325
                Fax: 580 221 7380
                Email: mjharrison@noble.org
                Insert Length: 784 Std Error: 0.00
                Plate: 114 row: D column: 02
                Seq primer: TCACACAGGAACACGCTATGAC.
FEATURES       Location/Qualifiers
                source
                1..784
                /organism="Medicago truncatula"
                /mol_type="mRNA"
                /db_xref="taxon:3880"
                /clone="NF114D02PL"
                /tissue_type="leaf"
                /dev_stage="trifoliolate"
                /clone_lib="Phosphate starved leaf"
                /note="Vector: Lambda Zap; At the trifoliolate stage, M.
                truncatula plants were transplanted to phosphate-free sand
                and grown for a further 30 days. During this 30 day
                period, the plants were fertilized twice weekly with 1/2
                Hoaglands solution containing only 20uM potassium
                phosphate. RNA was prepared from above ground tissues."
ORIGIN
Query Match      86.5%; Score 87.4; DB 5; Length 784;
Best Local Similarity 93.8%; Pred. No. 2.1e-13;
Matches 91; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CGGTCATCTTTGATTATAGGGATTTGGGGATTTGGCCCTATTGGTTAAAAAATG 60
    |||||
Db 194 CGGTCATCTTTGATTATAGGGATTTGGGGATTTGGCCCTATTGGTTAAAAAATG 135
    |||||

Qy 61 AGCTGATTTACAAAAATTTAACGCGAATTAAATTCG 97
    |||||
Db 134 AGCTGATTTACAAAAATTTAACGCGAATTCCTGCAG 98
    |||||

RESULT 12
BQ156587/c
LOCUS      BQ156587 785 bp mRNA linear EST 24-APR-2002
DEFINITION NF094D10IR1F090 Irradiated Medicago truncatula cDNA clone
            NF094D10IR 5', mRNA sequence.
ACCESSION  BQ156587
VERSION     BQ156587.1 GI:20293646
KEYWORDS    EST.
SOURCE      Medicago truncatula (barrel medic)
ORGANISM    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE    1 (bases 1 to 785)
AUTHORS      Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
            Flores,H.R., Inman,J.T., Weller,J.W. and May,G.D.
TITLE        Expressed Sequence Tags from the Samuel Roberts Noble Foundation
            Medicago truncatula irradiated library
JOURNAL      Unpublished (2001)

KEYWORDS      EST.
SOURCE         Medicago truncatula (barrel medic)
ORGANISM       Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
                Medicago.
REFERENCE      1 (bases 1 to 796)
AUTHORS        Watson,B.S., Shin,H.-S., Lopez-Meyer,M., Scott,A.D., Harris,A.R.,
                Gonzales,R.A., Bell,C.J., Inman,J.T., Waugh,M.E., Sullivan,J.P.,
                May,G.D. and Paiva,N.L.
TITLE          Expressed Sequence Tags from the Samuel Roberts Noble Foundation
                Medicago truncatula Phoma-infected library
JOURNAL        Unpublished (2002)
COMMENT        Contact: Paiva NL
                Plant Biology Division
                The Samuel Roberts Noble Foundation
                2510 Sam Noble Parkway, Ardmore, OK 73402, USA

COMMENT        Contact: May GD
                Plant Biology Division
                The Samuel Roberts Noble Foundation
                2510 Sam Noble Parkway, Ardmore, OK 73402, USA
                Tel: 580 224 6650
                Fax: 580 224 6692
                Email: gdmay@noble.org
                Insert Length: 785 Std Error: 0.00
                Plate: 094 row: D column: 10
                Seq primer: TCACACAGGAACACGCTATGAC.
FEATURES       Location/Qualifiers
                source
                1..785
                /organism="Medicago truncatula"
                /mol_type="mRNA"
                /db_xref="taxon:3880"
                /clone="NF094D10IR"
                /tissue_type="seedlings"
                /dev_stage="seedling"
                /clone_lib="Irradiated"
                /note="Vector: Lambda Zap; Seedlings were exposed either
                to 100 Gy gamma or 0.5, 1, 5, or 10 kJ/m2 UV irradiation.
                Gamma-irradiated samples were harvested at 6, 12, 24 and
                48 hours after treatment. UV-irradiated samples were
                harvested 24 hours post-treatment. cDNA was prepared from
                polyA+ enriched, pooled samples of equivalent amounts of
                total RNA from each sample. The cDNA was directionally
                ligated into the Uni-Zap XR vector (Stratagene) and
                packaged using the Gigapack III Gold packaging extracts.
                Phagemids containing cDNA inserts were in vivo excised
                from the recombinant Uni-Zap XR vector using ExAssist
                helper phage and the E. coli strain XL1-Blue MRF'
                (Stratagene). Excised plasmids were plated using SOLR
                cells."
ORIGIN
Query Match      86.5%; Score 87.4; DB 5; Length 785;
Best Local Similarity 93.8%; Pred. No. 2.1e-13;
Matches 91; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CGGTCATCTTTGATTATAGGGATTTGGGGATTTGGCCCTATTGGTTAAAAAATG 60
    |||||
Db 191 CGGTCATCTTTGATTATAGGGATTTGGGGATTTGGCCCTATTGGTTAAAAAATG 132
    |||||

Qy 61 AGCTGATTTACAAAAATTTAACGCGAATTAAATTCG 97
    |||||
Db 131 AGCTGATTTACAAAAATTTAACGCGAATTCCTGCAG 95
    |||||

RESULT 13
BQ139350/c
LOCUS      BQ139350 796 bp mRNA linear EST 26-APR-2002
DEFINITION NF014C09PH1F1070 Phoma-infected Medicago truncatula cDNA clone
            NF014C09PH 5', mRNA sequence.
ACCESSION  BQ139350
VERSION     BQ139350.1 GI:20275476
KEYWORDS    EST.
SOURCE      Medicago truncatula (barrel medic)
ORGANISM    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE    1 (bases 1 to 796)
AUTHORS      Watson,B.S., Shin,H.-S., Lopez-Meyer,M., Scott,A.D., Harris,A.R.,
                Gonzales,R.A., Bell,C.J., Inman,J.T., Waugh,M.E., Sullivan,J.P.,
                May,G.D. and Paiva,N.L.
TITLE          Expressed Sequence Tags from the Samuel Roberts Noble Foundation
                Medicago truncatula Phoma-infected library
JOURNAL        Unpublished (2002)
COMMENT        Contact: Paiva NL
                Plant Biology Division
                The Samuel Roberts Noble Foundation
                2510 Sam Noble Parkway, Ardmore, OK 73402, USA

```

Tel: 580 221 7317
Fax: 580 221 7380
Email: nipaiva@noble.org
Insert Length: 796 Std Error: 0.00
Plate: 014 row: C column: 09
Seq primer: TCACACGGAACAGCTATGAC.

FEATURES

source

Location/Qualifiers
1. .796
/organism="Medicago truncatula"
/mol_type="mRNA"
/db_xref="taxon:3880"
/clone="NF014C09PH"
/tissue_type="leaf"

/dev_stage="Pathogen-induced, young trifoliolate"
/note="Vector: pBluescript SK(-); Young trifoliolate leaves of Medicago truncatula were excised and dip-inoculated in humid dishes. Pools of leaves were harvested at 0, 15, and 30 minutes and 1, 2, 3, 6, 14, 24, 48, 72, and 96, hours, and used to prepare total RNA. cDNA was prepared from polyA+ enriched, pooled samples of equivalent amounts of total RNA from each sample. The cDNA was directionally ligated into the Uni-Zap XR vector (Stratagene) and packaged using the Gigapack III Gold packaging extracts. Phagemids containing cDNA inserts were in vivo excised from the recombinant Uni-Zap XR vector using ExAssist helper phage and the E. coli strain XL1-Blue MRF' (Stratagene). Excised plasmids were plated using SOLR cells."

ORIGIN

Query Match 86.5%; Score 87.4; DB 5; Length 796;
Best Local Similarity 93.8%; Pred. No. 2.1e-13;
Matches 91; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CGGTCTATCTTTTGATTATAAGGGATTTGGGGATTTCCGCCCTATTGGTTAAAAATG 60
Db 191 CGGTCTATCTTTTGATTATAAGGGATTTGGGGATTTCCGCCCTATTGGTTAAAAATG 132

Qy 61 AGCTGATTTACAAAAATTTAACGCGAATTAATCTG 97
Db 131 AGCTGATTTACAAAAATTTAACGCGAATTCCTGCAG 95

RESULT 14
BQ153647/c

LOCUS

DEFINITION NF040H05IR1F1047 Irradiated Medicago truncatula cDNA clone

ACCESSION BQ153647

VERSION

KEYWORDS BQ153647.1 GI:20290706

SOURCE

ORGANISM Medicago truncatula (barrel medic)

REFERENCE

1 (bases 1 to 796)
Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J., Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.
Expressed Sequence Tags from the Samuel Roberts Noble Foundation

JOURNAL

COMMENT

Medicago truncatula irradiated library
Unpublished (2001)
Contact: May GD
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 224 6650
Fax: 580 224 6692
Email: gdmay@noble.org
Insert Length: 796 Std Error: 0.00

Plate: 040 row: H column: 05
Seq primer: TCACACGGAACAGCTATGAC.

FEATURES

source

Location/Qualifiers
1. .796
/organism="Medicago truncatula"
/mol_type="mRNA"
/db_xref="taxon:3880"
/clone="NF040H05IR"
/tissue_type="seedling"
/dev_stage="seedling"
/clone_lib="Irradiated"
/note="Vector: Lambda Zap; Seedlings were exposed either to 100 Gy gamma or 0.5, 1, 5, or 10 kJ/m2 UV irradiation. Gamma-irradiated samples were harvested at 6, 12, 24 and 48 hours after treatment. UV-irradiated samples were harvested 24 hours post-treatment. cDNA was prepared from polyA+ enriched, pooled samples of equivalent amounts of total RNA from each sample. The cDNA was directionally ligated into the Uni-Zap XR vector (Stratagene) and packaged using the Gigapack III Gold packaging extracts. Phagemids containing cDNA inserts were in vivo excised from the recombinant Uni-Zap XR vector using ExAssist helper phage and the E. coli strain XL1-Blue MRF' (Stratagene). Excised plasmids were plated using SOLR cells."

ORIGIN

Query Match 86.5%; Score 87.4; DB 5; Length 796;
Best Local Similarity 93.8%; Pred. No. 2.1e-13;
Matches 91; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CGGTCTATCTTTTGATTATAAGGGATTTGGGGATTTCCGCCCTATTGGTTAAAAATG 60
Db 193 CGGTCTATCTTTTGATTATAAGGGATTTGGGGATTTCCGCCCTATTGGTTAAAAATG 134

Qy 61 AGCTGATTTACAAAAATTTAACGCGAATTAATCTG 97
Db 133 AGCTGATTTACAAAAATTTAACGCGAATTCCTGCAG 97

RESULT 15

BQ154182/c

LOCUS

DEFINITION NF056D06IR1F1057 Irradiated Medicago truncatula cDNA clone

ACCESSION BQ154182

VERSION

KEYWORDS BQ154182.1 GI:20291241

SOURCE

ORGANISM Medicago truncatula (barrel medic)

REFERENCE

1 (bases 1 to 798)
Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J., Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.
Expressed Sequence Tags from the Samuel Roberts Noble Foundation

JOURNAL

COMMENT

Medicago truncatula irradiated library
Unpublished (2001)
Contact: May GD
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 224 6650
Fax: 580 224 6692
Email: gdmay@noble.org
Insert Length: 798 Std Error: 0.00
Plate: 056 row: D column: 06

JOURNAL

COMMENT

Medicago truncatula irradiated library
Unpublished (2001)
Contact: May GD
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 224 6650
Fax: 580 224 6692
Email: gdmay@noble.org
Insert Length: 798 Std Error: 0.00
Plate: 056 row: D column: 06
Seq primer: TCACACGGAACAGCTATGAC.

FEATURES

source

Location/Qualifiers
1. .798
/organism="Medicago truncatula"

```
/mol_type="mRNA"  
/db_xref="taxon:3880"  
/clone="NF056D061R"  
/tissue_type="seedlings"  
/dev_stage="seedling"  
/clone_lib="Irradiated"  
/note="Vector: Lambda Zap; Seedlings were exposed either  
to 100 Gy gamma or 0.5, 1, 5, or 10 kJ/m2 UV irradiation.  
Gamma-irradiated samples were harvested at 6, 12, 24 and  
48 hours after treatment. UV-irradiated samples were  
harvested 24 hours post-treatment. cDNA was prepared from  
polyA+ enriched, pooled samples of equivalent amounts of  
total RNA from each sample. The cDNA was directionally  
ligated into the Uni-Zap XR vector (Stratagene) and  
packaged using the Gigapack III Gold packaging extracts.  
Phagemids containing cDNA inserts were in vivo excised  
from the recombinant Uni-Zap XR vector using ExAssist  
helper phage and the E. coli strain XL1-Blue MRF'  
(Stratagene). Excised plasmids were plated using SOLR  
cells."
```

ORIGIN

```
Query Match      86.5%; Score 87.4; DB 5; Length 798;  
Best Local Similarity 93.8%; Pred. No. 2.1e-13;  
Matches 91; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
  
Qy      1  CGGTCTATCTTTGATTATTAAGGGATTTTGGGGATTTTCGGCCTATTGGTTAAAAATG 60  
        |||||||  
Db      191 CGGTCTATCTTTGATTATTAAGGGATTTTGGGGATTTTCGGCCTATTGGTTAAAAATG 132  
        |||||||  
  
Qy      61  AGCTGATTTTACAAAAATTTTAAACGCGAATTAAATTCG 97  
        |||||||  
Db      131 AGCTGATTTTACAAAAATTTTAAACGCGAATTCCTCGAG 95  
        |||||||
```

Search completed: July 14, 2005, 23:22:30
Job time : 967.667 secs

Result No.	Score	Query		DB	ID	Description
		Match	Length			
C 1	101	100.0	142	6	AR356490	Sequence
C 2	101	100.0	142	6	AR538046	Sequence
C 3	101	100.0	228	6	E00019	DNA coding
C 4	101	100.0	240	1	PMOENDO	M10199 Plasmid pMW
C 5	101	100.0	251	6	E00018	DNA coding
C 6	101	100.0	251	6	I01644	Sequence 1
C 7	101	100.0	344	11	HUMUT5345	L18624 Human chr10
C 8	101	100.0	400	6	BD195256	Nucleotide
C 9	101	100.0	456	6	E00892	Synthetic D
C 10	101	100.0	456	6	E01156	DNA fragment
C 11	101	100.0	456	6	E01274	DNA encoding
C 12	101	100.0	456	6	E01302	DNA encoding
C 13	101	100.0	466	6	AX260098	Sequence
C 14	101	100.0	573	6	AX260150	Sequence
C 15	101	100.0	693	6	A43586	Sequence 11
C 16	101	100.0	693	6	AR116755	Sequence
C 17	101	100.0	998	1	AY559171	Pseudomon
C 18	101	100.0	1011	1	SMTEAAGE	X97254 S.marcescen
C 19	101	100.0	1012	2	CEC11F10	Z92776 Caenorhabdi

```

ORGANISM Unknown.
REFERENCE 1 (bases 1 to 142)
AUTHORS Kunsch,C.A., Choi,G.A., Barash,S.C., Dillon,P.J., Fannon,M.R. and Rosen,C.A.
TITLE Staphylococcus aureus polynucleotides and sequences
JOURNAL Patent: US 6737248-A 2608 18-MAY-2004;
FEATURES Location/Qualifiers
    source 1..142
        /organism="unknown"
        /mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 142;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAAATAG 60
Db 107 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAAATAG 48

Qy 61 GGGTTCGCGCACATTTCCCGGAAAAGTGCCACCTGAGTC 101
Db 47 GGGTTCGCGCACATTTCCCGGAAAAGTGCCACCTGAGTC 7

RESULT 3
E00019/c
LOCUS DNA coding for Escherichia coli penicillinase.
DEFINITION E00019
ACCESSION E00019
VERSION E00019.1 GI:2168327
KEYWORDS JP 1981154999-A/2.
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE 1 (bases 1 to 228)
AUTHORS Uotutaa,G. and Karen,T.
TITLE SYNTHESIS OF SELECTIVE AGED PROTEIN OR POLYPEPTIDE IN BACTERIA
JOURNAL Patent: JP 1981154999-A 2 30-NOV-1981;
COMMENT OS Escherichia coli
PN JP 1981154999-A/2
PD 30-NOV-1981
PF 09-APR-1981 JP 1981052488
PR 11-APR-1980 US 80 139225
PI UOJUTAA GIRUBAATO, KAREN TARUMATSUJI
PC C12P21/00,C07H21/00,C12N1/00,C12N15/00//C12R1/19; CC
strandedness: Double;
CC topology: Linear;
CC anti-sense: No;
CC *source: clone=pKT218;
FH Key Location/Qualifiers
FH CDS 210..>228
FT /product='E.coli penicillinase'.
FEATURES
    source 1..228
        Location/Qualifiers
        1..228
            /organism="Escherichia coli"
            /mol_type="genomic DNA"
            /db_xref="taxon:562"

ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 228;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAAATAG 60
Db 175 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAAATAG 116

Qy 61 GGGTTCGCGCACATTTCCCGGAAAAGTGCCACCTGAGTC 101

```

```

Db 115 GGGTTCGCGCACATTTCCCGGAAAAGTGCCACCTGAGTC 75

RESULT 4
PMOENDO/c
LOCUS DNA 240 bp linear BCT 26-APR-1993
DEFINITION Plasmid pMM110 region of endo VII cleavage sites near cruciform structures.
ACCESSION M10199
VERSION M10199.1 GI:150826
KEYWORDS
SOURCE Plasmid pMM110
ORGANISM Plasmid pMM110
other sequences; plasmids.
REFERENCE 1 (bases 1 to 240)
AUTHORS Kemper,B., Jensch,F., von Depka-Prondzynski,M., Fritz,H.J., Borgmeyer,U. and Mizunuchi,K.
TITLE Resolution of Holliday structures by endonuclease VII as observed in interactions with cruciform DNA
JOURNAL Cold Spring Harb. Symp. Quant. Biol. 49, 815-825 (1984)
MEDLINE 85153063
PUBMED 6397324
COMMENT Original source text: Plasmid pMM110 DNA.
FEATURES
    source 1..240
        Location/Qualifiers
        1..240
            /organism="Plasmid pMM110"
            /mol_type="genomic DNA"
            /db_xref="taxon:2599"
            /plasmid="Plasmid pMM110"
            Unreported.

ORIGIN
Query Match 100.0%; Score 101; DB 1; Length 240;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAAATAG 60
Db 151 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAAATAG 92

Qy 61 GGGTTCGCGCACATTTCCCGGAAAAGTGCCACCTGAGTC 101
Db 91 GGGTTCGCGCACATTTCCCGGAAAAGTGCCACCTGAGTC 51

RESULT 5
E00018/c
LOCUS DNA coding for Escherichia coli penicillinase.
DEFINITION E00018
ACCESSION E00018
VERSION E00018.1 GI:2168326
KEYWORDS JP 1981154999-A/1.
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE 1 (bases 1 to 251)
AUTHORS Uotutaa,G. and Karen,T.
TITLE SYNTHESIS OF SELECTIVE AGED PROTEIN OR POLYPEPTIDE IN BACTERIA
JOURNAL Patent: JP 1981154999-A 1 30-NOV-1981;
COMMENT OS Escherichia coli
PN JP 1981154999-A/1
PD 30-NOV-1981
PF 09-APR-1981 JP 1981052488
PR 11-APR-1980 US 80 139225
PI UOJUTAA GIRUBAATO, KAREN TARUMATSUJI
PC C12P21/00,C07H21/00,C12N1/00,C12N15/00//C12R1/19; CC
strandedness: Double;
CC topology: Linear;
CC anti-sense: No;
CC fragment type: N-Terminal Fragment;
CC *source: clone=pKT241;

```

REFERENCE AUTHORS

```

FH Key Location/Qualifiers
FT source 1..400
FT /organism='Unidentified'.
FEATURES
  source
    1..400
    Location/Qualifiers
    /organism='Unidentified'
    /mol_type='genomic DNA'
    /db_xref='taxon:32644'
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 400;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 60
Db 165 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 106
Qy 61 GGGTTCCGCGCACATTTCCCGGAAAAAGTGCACCTGACGTC 101
Db 105 GGGTTCCGCGCACATTTCCCGGAAAAAGTGCACCTGACGTC 65
RESULT 9
E00892/c
LOCUS E00892 456 bp DNA linear PAT 29-SEP-1997
DEFINITION Synthetic DNA encoding fused polypeptide between E coli
beta-lactamase and human beta-urogastrone.
ACCESSION E00892
VERSION E00892.1 GI:2169153
KEYWORDS JP 1986149089-A/1.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 456)
AUTHORS Okai.H., Momota.Y., Kumakura.T., Tochifusa.N., Kitazawa.T.,
Ojida.K. and Matsuehiro.A.
TITLE POLYPEPTIDE SECRETION DEVELOPMENT VECTOR, MICROORGANISM TRANSFORMED
WITH SAME, AND PRODUCTION OF POLYPEPTIDE WITH MICROORGANISM
JOURNAL Patent: JP 1986149089-A 1 07-JUL-1986;
EARTH CHEM CORP LTD
COMMENT OS Artificial gene
OC Artificial sequence; Genes.
PN JP 1986149089-A/1
PD 07-JUL-1986
PF 21-DEC-1984 JP 1984271206
PI OKAI HIDEO, MOMOTA YUTAKA, KUMAKURA TAKESHI,
PI TOCHIFUSA NORIYUKI,
PI KITAZAWA TOSHIKI, OJIDA KAZUHIRO, MATSUSHIRO AIZO PC
C12N15/00,C12N1/20,C12P21/00,(C12N1/20,C12R1:19),(C12P21/00,PC
C12R1:19);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: strain=HB101;
CC *source: clones=pVG201;
CC Feature is identified by experimental;
FH Key Location/Qualifiers
FH promoter 125..170
FH of beta-lactamase
FT RBS 200..203
FT CDS 209..438
FT /product='beta-urogastrone precursor' FT
FT sig_peptide 209..277
FT /product='signal peptide of beta-lactonase' FT
FT mat_peptide 278..435
FT /product='beta-urogastrone mature peptide'.
FT Location/Qualifiers
  1..456
  /organism='synthetic construct'
  /mol_type='genomic DNA'
FEATURES
  source
    1..456
    Location/Qualifiers
    /organism='synthetic construct'
    /mol_type='genomic DNA'

```

```

ORIGIN
/db_xref='taxon:32630'
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 60
Db 173 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 114
Qy 61 GGGTTCCGCGCACATTTCCCGGAAAAAGTGCACCTGACGTC 101
Db 113 GGGTTCCGCGCACATTTCCCGGAAAAAGTGCACCTGACGTC 73
RESULT 10
E01156/c
LOCUS E01156 456 bp DNA linear PAT 29-SEP-1997
DEFINITION DNA fragment which secretes beta urogastrone.
ACCESSION E01156
VERSION E01156.1 GI:2169415
KEYWORDS JP 1987083890-A/1.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 456)
AUTHORS Yoshikawa,K., Momota,Y., Kajifusa,N., Koide,T. and Okai,H.
TITLE POLYPEPTIDE SECRETION EXPRESSION VECTOR, MICROORGANISM TRANSFORMED
BY SAID VECTOR AND PRODUCTION OF POLYPEPTIDE USING SAID
JOURNAL Patent: JP 1987083890-A 1 17-APR-1987;
EARTH CHEM CORP LTD
COMMENT OS Artificial gene
OC Artificial sequence; Genes.
PN JP 1987083890-A/1
PD 17-APR-1987
PF 09-OCT-1985 JP 1985225393
PI YOSHIKAWA KAZUTOSHI, MOMOTA YUTAKA, KAJIFUSA NORIYUKI, PI
KOIDE TAKAO,
PI OKAI HIDEO
PC C12N15/00,C12N1/20,C12P21/00,(C12N1/20,C12R1:19),(C12N1/20,PC
C12R1:125);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: clone=pUG201;
CC Key Location/Qualifiers
FH promoter 125..170
FH /note='beta lactamase promoter' FT RBS
FT CDS 209..439
FT /product='beta urogastrone'
FT sig_peptide 209..277
FT mat_peptide 278..436
FT /product='beta urogastrone'.
FT Location/Qualifiers
  1..456
  /organism='synthetic construct'
  /mol_type='genomic DNA'
  /db_xref='taxon:32630'
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 60
Db 173 AGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 114
Qy 61 GGGTTCCGCGCACATTTCCCGGAAAAAGTGCACCTGACGTC 101

```

```
Db 113 GGGTTCCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 73
|||||
RESULT 11
E01274/c
LOCUS E01274 456 bp DNA linear PAT 29-SEP-1997
DEFINITION DNA encoding beta-urogastron fused with DNA encoding a promoter and
signal peptide of beta-lactamase.
ACCESSION E01274
VERSION E01274.1 GI:2169533
KEYWORDS JP 1987179398-A/1.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 456)
AUTHORS Okai,H., Kumakura,T., Kawamoto,S., Adachi,S., Matsubara,A.,
Ojida,K., Yano,M., Mihara,S., Matsushiro,A. and Yanaihara,N.
TITLE PRODUCTION OF BETA-UROGASTRONE
JOURNAL EARTH CHEM CORP LTD
COMMENT OS Artificial gene
OS Homo sapiens
PN JP 1987179398-A/1
PD 06-AUG-1987
PF 31-JAN-1986 JP 1986021032
PI OKAI HIDEO, KUMAKURA TAKESHI, KAWAMOTO SHOICHI, ADACHI SHOJI,
MATSUBARA AKIMASA, OJIDA KAZUHIDE, YANO MAKI, MIHARA SHIGERU,
MATSUSHIRO AIZO, YANAIHARA NOBORU
PC C12P21/00,C12N15/00,(C12P21/00,C12R1:91);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No; Location/Qualifiers
FH Key
FH Promoter 125..170
FT RBS 200..203
FT sig_peptide 209..277
FT mat_peptide 278..436
FT CDS 209..439
FT /product='beta-urogastron'
FT /product='beta-urogastron'
FEATURES
source
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGAATGTTATTAGAAAATAAACAAATAG 60
|||||
Db 173 AGGGTTATTGTCATGAGCGGATACATATTGAATGTTATTAGAAAATAAACAAATAG 114
|||||
Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101
|||||
Db 113 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 73
|||||
RESULT 13
AX260098/c
LOCUS AX260098 466 bp DNA linear PAT 26-OCT-2001
DEFINITION Sequence 60 from Patent WO0172774.
ACCESSION AX260098
VERSION AX260098.1 GI:16509129
KEYWORDS Drosophila melanogaster (fruit fly)
SOURCE Drosophila melanogaster
ORGANISM Drosophila melanogaster
REFERENCE 1
AUTHORS Deak,P., Glover,D.M. and Midgley,C.
TITLE Cell cycle progression proteins
JOURNAL Patent: WO 0172774-A 60 04-OCT-2001;
Cyclacel Limited (GB)
FEATURES
source
ORIGIN
Query Match 100.0%; Score 101; DB 6; Length 456;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGAATGTTATTAGAAAATAAACAAATAG 60
|||||
Db 173 AGGGTTATTGTCATGAGCGGATACATATTGAATGTTATTAGAAAATAAACAAATAG 114
|||||
Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101
|||||
Db 113 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 73
|||||
RESULT 12
E01302/c
LOCUS E01302 456 bp DNA linear PAT 29-SEP-1997
DEFINITION DNA encoding human beta-urogastrone fused with DNA encoding
promoter and signal peptide of beta-lactamase.
ACCESSION E01302
VERSION E01302.1 GI:2169561
KEYWORDS JP 1987190083-A/1.
SOURCE synthetic construct
```

```
ORIGIN
/db_xref="taxon:7227"

Query Match      100.0%; Score 101; DB 6; Length 466;
Best Local Similarity 100.0%; Pred. No. 8.7e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGATGATTTAGAAAAATAAACAAATAG 60
    |||||||
Db 280 AGGGTTATTGTCATGAGCGGATACATATTGTAATGATGATTTAGAAAAATAAACAAATAG 221
    |||||||

Qy 61 GGGTTCCGCGCACATTTCCCGAAAAAGTGCACCTGACGTC 101
    |||||||
Db 220 GGGTTCCGCGCACATTTCCCGAAAAAGTGCACCTGACGTC 180
    |||||||

RESULT 14
AX260150/c      AX260150      573 bp      DNA      linear      PAT 26-OCT-2001
LOCUS
DEFINITION      Sequence 112 from Patent WO0172774.
ACCESSION      AX260150
VERSION        AX260150.1 GI:16509172
KEYWORDS
SOURCE
ORGANISM        Drosophila melanogaster (fruit fly)
                Drosophila melanogaster
                Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
                Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
                Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
AUTHORS        Deak, P., Glover, D.M. and Midgley, C.
TITLE          Cell cycle progression proteins
JOURNAL        Patent: WO 0172774-A 112 04-OCT-2001;
                Cyclacel Limited (GB)
FEATURES
source
1..573
/organism="Drosophila melanogaster"
/mol_type="unassigned DNA"
/db_xref="taxon:7227"

ORIGIN

Query Match      100.0%; Score 101; DB 6; Length 573;
Best Local Similarity 100.0%; Pred. No. 8.8e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGATGATTTAGAAAAATAAACAAATAG 60
    |||||||
Db 355 AGGGTTATTGTCATGAGCGGATACATATTGTAATGATGATTTAGAAAAATAAACAAATAG 296
    |||||||

Qy 61 GGGTTCCGCGCACATTTCCCGAAAAAGTGCACCTGACGTC 101
    |||||||
Db 295 GGGTTCCGCGCACATTTCCCGAAAAAGTGCACCTGACGTC 255
    |||||||

RESULT 15
A43586
LOCUS
DEFINITION      Sequence 11 from Patent WO9507357.
ACCESSION      A43586
VERSION        A43586.1 GI:2298779
KEYWORDS
SOURCE
ORGANISM        Cuphea lanceolata
                Cuphea lanceolata
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                rosids; Myrtales; Lythraceae; Cuphea.
                1 (bases 1 to 693)
                Toepfer, R., Bautor, J., Bothmann, H., Filsak, E.,
                Hoerlcke-Grandpierre, C., Klein, B., Martini, N., Mueller, A.,
                Schulte, W., Voetz, M., Walek, J. and Schell, J.
REFERENCE
AUTHORS
PROMOTERS
TITLE          Patent: WO 9507357-A 11 16-MAR-1995;
JOURNAL        MAX PLANCK GESELLSCHAFT (DE)
COMMENT        Other publication CA 2169093 950316
```

```
FEATURES
source
Other publication AU 7615494 950327.
Location/Qualifiers
1..693
/organism="Cuphea lanceolata"
/mol_type="unassigned DNA"
/db_xref="taxon:3930"
/clone="CLKASIG8"
/clone_lib="Genomic Lambda Fix II"

ORIGIN

Query Match      100.0%; Score 101; DB 6; Length 693;
Best Local Similarity 100.0%; Pred. No. 8.8e-20;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGATGATTTAGAAAAATAAACAAATAG 60
    |||||||
Db 592 AGGGTTATTGTCATGAGCGGATACATATTGTAATGATGATTTAGAAAAATAAACAAATAG 651
    |||||||

Qy 61 GGGTTCCGCGCACATTTCCCGAAAAAGTGCACCTGACGTC 101
    |||||||
Db 652 GGGTTCCGCGCACATTTCCCGAAAAAGTGCACCTGACGTC 692
    |||||||

Search completed: July 14, 2005, 14:03:24
Job time : 759.618 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 04:35:42 ; Search time 142.398 Seconds
(without alignments)
4198.742 Million cell updates/sec

Title: US-09-482-682-8_COPY_7369_7469

Perfect score: 101
Sequence: 1 agggttattgtctcatgagc.....gaaagtgccacctgacgtc 101

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04.*

- 1: geneseqn1980s.*
- 2: geneseqn1990s.*
- 3: geneseqn2000s.*
- 4: geneseqn2001as.*
- 5: geneseqn2001bs.*
- 6: geneseqn2002as.*
- 7: geneseqn2002bs.*
- 8: geneseqn2003as.*
- 9: geneseqn2003bs.*
- 10: geneseqn2003cs.*
- 11: geneseqn2003ds.*
- 12: geneseqn2004as.*
- 13: geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	101	100.0	142	2	AAV76919
C 2	101	100.0	228	1	AAN10032
C 3	101	100.0	251	1	AAN10031
C 4	101	100.0	400	2	AAV31229 E. coli J
C 5	101	100.0	456	1	AAN60624 Plasmid p
C 6	101	100.0	456	1	AAN71080 Sequence
C 7	101	100.0	456	1	AAN70833 Beta-urog
C 8	101	100.0	456	1	AAN81765 Sequence
C 9	101	100.0	466	6	ABA90413 Drosophil
C 10	101	100.0	487	2	AAx21173 Polynucle
C 11	101	100.0	535	2	AAx21149 Polynucle
C 12	101	100.0	573	6	ABA90456 Drosophil
C 13	101	100.0	605	12	ADH58311 Electroph
C 14	101	100.0	776	4	AAx30560 DNA encod
C 15	101	100.0	776	4	AAx27819
C 16	101	100.0	776	4	ABK42984 Genomic s
C 17	101	100.0	776	4	AAAL07344 Human rep
C 18	101	100.0	776	4	AAAL03229 Human rep
C 19	101	100.0	776	4	AAAL06588 Human rep
C 20	101	100.0	776	4	AAAL07340 Human rep

C 21	101	100.0	776	5	ABA14573
C 22	101	100.0	776	5	AAS34681
C 23	101	100.0	776	8	ADA41574
C 24	101	100.0	776	8	ACC50905
C 25	101	100.0	776	8	ABZ71508
C 26	101	100.0	776	9	ADB91869
C 27	101	100.0	776	9	ADB61140
C 28	101	100.0	776	10	ADB94622
C 29	101	100.0	776	10	ADC74663
C 30	101	100.0	776	10	ADA57709
C 31	101	100.0	776	12	ADN41551
C 32	101	100.0	845	4	AAS30559
C 33	101	100.0	845	4	AAS27818
C 34	101	100.0	845	4	ABK42983
C 35	101	100.0	845	4	AAS41807
C 36	101	100.0	845	4	AAS41855
C 37	101	100.0	845	4	AAK85485
C 38	101	100.0	845	4	AAK85434
C 39	101	100.0	845	4	AAAL07343
C 40	101	100.0	845	4	AAAL06587
C 41	101	100.0	845	4	AAAL07339
C 42	101	100.0	845	4	AAAL03228
C 43	101	100.0	845	5	ABA14572
C 44	101	100.0	845	5	AAS34680
C 45	101	100.0	845	9	ADB61139

ALIGNMENTS

RESULT 1

AAV76919/c
ID AAV76919 standard; DNA; 142 BP.
XX
AC AAV76919;
XX
DT 16-MAR-1999 (first entry)
XX

DE Staphylococcus aureus contig SEQ ID #2608.

Computer readable medium; vaccine; S.aureus infection; immunodetection;
cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
skin infection; surgical wound infection; scalded skin syndrome;
toxic shock syndrome; ds.

OS Staphylococcus aureus.

XX EP786519-A2.

XX 30-JUL-1997.

XX 07-JAN-1997; 97EP-00100117.

XX 05-JAN-1996; 96US-0009861P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Kunsch CA, Choi GH, Barash SC, Dillon PJ, Fannon MR, Rosen CA;

XX WPI; 1997-374922/35.

XX Polynucleotide(s) and proteins derived from Staphylococcus aureus -
stored on computer readable medium and used in the production of anti-
S.aureus vaccines.

XX Claim 1; Page 2287; 3271pp; English.

XX This sequence represents one of 5191 Staphylococcus aureus DNA sequences
of the invention. The DNA sequences are recorded on a computer readable
medium, preferably selected from a floppy or hard disk, random access
memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
the S.aureus DNA sequences allows putative functions to be assigned so
that protein-encoding or regulatory regions of commercial, therapeutic or

CC industrial importance can be obtained. Specifically, sequences which are
 CC likely to encode antigens have been identified and these polypeptides can
 CC be used in a vaccine composition against *S.aureus* infection. The
 CC polypeptides can also be used in a kit for the immunodetection of
 CC *S.aureus* in a sample. *S.aureus* is implicated in numerous human diseases,
 CC including cellulitis, eyelid infections, food poisoning, osteomyelitis,
 CC skin and surgical wound infections, scalded skin syndrome, toxic shock
 CC syndrome, etc. Organisms transformed with the DNA sequences can be used
 CC for recombinant production of the polypeptides. The new DNA sequences
 CC (and their fragments) are useful as primers or probes for isolating
 CC homologues of any of the *S.aureus* DNA sequences contained on the computer
 CC readable medium

XX SQ Sequence 142 BP; 45 A; 25 C; 26 G; 45 T; 0 U; 1 Other;

Query Match 100.0%; Score 101; DB 2; Length 142;
 Best Local Similarity 100.0%; Pred. No. 2.1e-21;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAAATAG 60
 |||||
 Db 107 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAAATAG 48
 |||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101
 |||||
 Db 47 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 7
 |||||

RESULT 2

AAAN10032/c
 ID AAAN10032 standard; DNA; 228 BP.

XX AC AAAN10032;

XX DT 13-AUG-1992 (first entry)

XX DE Sequence of the pKT218 EcoRI-PstI penicillinase gene fragment.

XX KW Cloning vehicle; bacterial vector; transformed host; penicillinase;
 XX insulin; ds.

XX OS Escherichia coli.

XX FH	Key	Location/Qualifiers
XX FT	misc_feature	1..4
XX FT		/*tag= a
XX FT		/label= sticky end
XX FT	misc_feature	225..228
XX FT		/*tag= b
XX FT		/label= sticky end

XX EP38182-A.

XX PN 21-OCT-1981.

XX PF 09-APR-1981; 81EP-00301561.

XX PR 11-APR-1980; 80US-00139225.

XX PA (HARD) HARVARD COLLEGE.

XX PI Gilbert W, Talmadge K;

XX DR WPI; 1981-80125D/44.

XX DR P-PSDB; AAP10039.

XX PS Synthesis of mature protein or polypeptide - by using bacterial host
 XX transformed by cloned vehicle contg. DNA fragment etc.

XX Example; Fig 3; 34pp; English.

XX CC The closest identifiable promoter for the penicillinase gene in pKT241
 CC (AAAN10031) is located in the region 14 to 20 nucleotides before its

CC translational start signal. In the examples, the 3' end of pKT241 was
 CC attached to the signal DNA sequence of the DNA fragment (19) for rat
 CC preproinsulin (see AAAN10033). The closest identifiable promoter for the
 CC penicillinase gene in pKT218 (AAAN10032) is located in the region 14 to 20
 CC nucleotides before its translational start signal. In the examples, the
 CC 3' end of pKT218 was attached to the signal DNA sequence of the DNA
 CC fragment (CB6) for rat preproinsulin (see AAAN10034)

XX SQ Sequence 228 BP; 72 A; 37 C; 46 G; 73 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 1; Length 228;
 Best Local Similarity 100.0%; Pred. No. 2.3e-21;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAAATAG 60
 |||||
 Db 175 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTTAGAAAAATAAACAAATAG 116
 |||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101
 |||||
 Db 115 GGGTTCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 75
 |||||

RESULT 3

AAAN10031/c
 ID AAAN10031 standard; DNA; 251 BP.

XX AC AAAN10031;

XX DT 13-AUG-1992 (first entry)

XX DE Sequence of the pKT241 EcoRI-PstI penicillinase gene fragment.

XX KW Cloning vehicle; bacterial vector; transformed host; penicillinase;
 XX insulin; ds.

XX OS Escherichia coli.

XX FH	Key	Location/Qualifiers
XX FT	misc_feature	1..4
XX FT		/*tag= a
XX FT		/label= sticky end
XX FT	misc_feature	248..251
XX FT		/*tag= b
XX FT		/label= sticky end

XX EP38182-A.

XX PD 21-OCT-1981.

XX PF 09-APR-1981; 81EP-00301561.

XX PR 11-APR-1980; 80US-00139225.

XX PA (HARD) HARVARD COLLEGE.

XX PI Gilbert W, Talmadge K;

XX DR WPI; 1981-80125D/44.

XX DR P-PSDB; AAP10038.

XX PS Synthesis of mature protein or polypeptide - by using bacterial host
 XX transformed by cloned vehicle contg. DNA fragment etc.

XX Example; Fig 2; 34pp; English.

XX CC The closest identifiable promoter for the penicillinase gene in pKT241
 CC (AAAN10031) is located in the region 14 to 20 nucleotides before its
 CC translational start signal. In the examples, the 3' end of pKT241 was
 CC attached to the signal DNA sequence of the DNA fragment (19) for rat
 CC preproinsulin (see AAAN10033). The closest identifiable promoter for the
 CC penicillinase gene in pKT218 (AAAN10032) is located in the region 14 to 20
 CC nucleotides before its translational start signal. In the examples, the

CC 3' end of pKT218 was attached to the signal DNA sequence of the DNA
CC fragment (CB6) for rat preproinsulin (see AAN10034)

xx Sequence 251 BP; 74 A; 45 C; 50 G; 82 T; 0 U; 0 Other;
SQ Query Match 100.0%; Score 101; DB 1; Length 251;
Best Local Similarity 100.0%; Pred. No. 2.3e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGCGGATACATATTTGAATGTATTAGAAAAATAACAAATAG 60
Db |||||||
Qy 61 GGGTTCGGCGACATTTCCCGAAAAGTGCCACCTGAGTC 101
Db |||||||
115 GGGTTCGGCGACATTTCCCGAAAAGTGCCACCTGAGTC 75

RESULT 4

AAV31229/c
ID AAV31229 standard; DNA; 400 BP.

AC AAV31229;

DT 01-OCT-1998 (first entry)

xx E. coli J96 pathogenicity island contig #43.

DE PAI; pathogenicity island; uropathogenic E. coli detection; PAI IV; pHER;
KW PAI V; phev; vaccine; protective immune response; ds.

xx Escherichia coli.

xx WO9822575-A2.

xx 28-MAY-1998.

xx 21-NOV-1997; 97WO-US021347.

xx 22-NOV-1996; 96US-0031626P.

PR 14-OCT-1997; 97US-0061953P.

xx (HUMA-) HUMAN GENOME SCI INC.

PA (UTWI-) UNIV WISCONSIN.

xx Dillon PJ, Choi GH, Welch RA;

xx WPI; 1998-312461/27.

DR New isolated uropathogenic E. coli nucleotide sequences - used to develop
xx products for the detection of pathogenic E. coli and to elicit an immune
xx response to pathogenic E. coli.

PS Claim 21; Page 140-141; 250pp; English.

xx This sequence represents a E. coli strain J96 contig containing
xx pathogenicity island (PAI) sequences, and represents a nucleic acid
xx molecule of the invention. PAIs are large fragments of DNA which comprise
xx pathogenicity determinants. The sequences of the invention are taken from
xx PAI IV and PAI V. PAI IV is located at approximately 64 min (near phev)
xx on the E. coli chromosome and is greater than 170 kb. PAI V is located at
xx approximately 94 min (at pHER) on the E. coli chromosome and is
xx approximately 160 kb in size. Antibodies specific to the proteins encoded
xx by the PAI open reading frames of the invention can be used in kits to
xx detect uropathogenic E. coli. The proteins are used in vaccines to elicit
xx a protective immune response in an animal to the uropathogenic E. coli
xx strain J96

xx Sequence 400 BP; 106 A; 77 C; 91 G; 126 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 2; Length 400;
Best Local Similarity 100.0%; Pred. No. 2.5e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGCGGATACATATTTGAATGTATTAGAAAAATAACAAATAG 60
Db |||||||
Qy 61 GGGTTCGGCGACATTTCCCGAAAAGTGCCACCTGAGTC 101
Db |||||||
105 GGGTTCGGCGACATTTCCCGAAAAGTGCCACCTGAGTC 65

RESULT 5

AAN60624/c
ID AAN60624 standard; DNA; 456 BP.

xx AAN60624;

xx 25-MAR-2003 (revised)
DT 29-OCT-1991 (first entry)

DE Plasmid pUG201 sequence encoding beta-urogastrone.

xx Beta-lactamase signal peptide; pGH54; pGH55; ss.

xx Synthetic.

xx Key Location/Qualifiers

FH 125..170

FT promoter /*tag= a

FT RBS 200..203

FT /*tag= b

FT CDS 209..439

FT /*tag= c

FT sig_peptide 209..277

FT /*tag= d

FT /label= Beta-lactamase signal peptide

FT 278..436

FT /*tag= e

FT /label= Beta-urogastrone

xx WO8603779-A.

xx 03-JUL-1986.

xx 19-DEC-1985; 85WO-JP000696.

xx 21-DEC-1984; 84JP-00271206.

xx (EART) EARTH CHEM CO LTD.

PA (OHGA/) OHGAI H.

xx Ohgai H, Momota H, Kuwakura T, Kajifusa N, Kitazawa T, Oshiden K;

xx WPI; 1986-182911/28.

DR P-ESDB; AAP60678.

xx Recombinant vector for polypeptide secretion - contains signal peptide
xx sequence directly bonded to peptide-coding sequence.

PS Disclosure; Table 4; 79pp; Japanese.

xx The plasmid produces secreted beta-urogastrone in a transformed
xx expression system. Similar plasmids may be constructed where the
xx secretion signal may be coupled with eg. somatostatin, insulin, growth
xx hormone, interferon, IL-2, gastric inhibitory peptide, influenza B SA,
xx epidermal growth factor and thymosine-beta4. (Updated on 25-MAR-2003 to
xx correct PA field.)

xx Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;

Query Match 100.0%; Score 101; DB 1; Length 456;
Best Local Similarity 100.0%; Pred. No. 2.6e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGATTATTAGAAAAATAAACAATAG 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
173 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGATTATTAGAAAAATAAACAATAG 114

Qy 61 GGGTTCCGCGCACATTTCCCGGAAAAAGTGCACCTGACGTC 101
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
113 GGGTTCCGCGCACATTTCCCGGAAAAAGTGCACCTGACGTC 73

RESULT 6
AAN71080/C
ID AAN71080 standard; DNA; 456 BP.
XX
AC AAN71080;
XX
XX 25-MAR-2003 (revised)
DT 10-MAR-2003 (revised)
DT 13-MAY-1991 (first entry)
XX
XX Sequence encoding beta-urogastrone.
DE
XX pUGT 150s; beta-UG; ds.
XX
XX Escherichia coli.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH promoter 125..170
FT /*tag= a
FT CDS 209..439
FT /*tag= b
FT /transl_except= (pos:434..436,aa:Arg)
XX
XX JP62190083-A.
XX
XX 20-AUG-1987.
XX
XX 14-FEB-1986; 86JP-00031415.
XX
XX 14-FEB-1986; 86JP-00031415.
XX
XX (EART ) EARTH SEIYAKU KK.
XX
XX WPI; 1987-273761/39.
XX
XX Expression vector contg. multiple information units is used to transform
PT host all for increased prodn. of polypeptides.
XX
XX Disclosure; Page 553; 34pp; Japanese.
XX
XX Sequence encodes beta-urogastrone under the control of a tac promoter.
CC The peptide may be expressed from plasmid pUGT 150s in a transformed
CC E.coli host. The plasmid may carry several separately expressing
CC sequences comprising a tac promoter, SD site, signal peptide, and coding
CC sequence, to produce beta-UG in high yield. (Updated on 10-MAR-2003 to
CC add missing OS field.) (Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 101; DB 1; Length 456;
Best Local Similarity 100.0%; Pred. No. 2.6e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGATTATTAGAAAAATAAACAATAG 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
173 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGATTATTAGAAAAATAAACAATAG 114

Qy 61 GGGTTCCGCGCACATTTCCCGGAAAAAGTGCACCTGACGTC 101
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
113 GGGTTCCGCGCACATTTCCCGGAAAAAGTGCACCTGACGTC 73

RESULT 7
AAN71080/C
ID AAN71080 standard; DNA; 456 BP.
XX
AC AAN71080;
XX
XX 25-MAR-2003 (revised)
DT 10-MAR-2003 (revised)
DT 13-MAY-1991 (first entry)
XX
XX Sequence encoding beta-urogastrone.
DE
XX pUGT 150s; beta-UG; ds.
XX
XX Escherichia coli.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH promoter 125..170
FT /*tag= a
FT CDS 209..439
FT /*tag= b
FT /transl_except= (pos:434..436,aa:Arg)
XX
XX JP62190083-A.
XX
XX 20-AUG-1987.
XX
XX 14-FEB-1986; 86JP-00031415.
XX
XX 14-FEB-1986; 86JP-00031415.
XX
XX (EART ) EARTH SEIYAKU KK.
XX
XX WPI; 1987-273761/39.
XX
XX Expression vector contg. multiple information units is used to transform
PT host all for increased prodn. of polypeptides.
XX
XX Disclosure; Page 553; 34pp; Japanese.
XX
XX Sequence encodes beta-urogastrone under the control of a tac promoter.
CC The peptide may be expressed from plasmid pUGT 150s in a transformed
CC E.coli host. The plasmid may carry several separately expressing
CC sequences comprising a tac promoter, SD site, signal peptide, and coding
CC sequence, to produce beta-UG in high yield. (Updated on 10-MAR-2003 to
CC add missing OS field.) (Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 101; DB 1; Length 456;
Best Local Similarity 100.0%; Pred. No. 2.6e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGATTATTAGAAAAATAAACAATAG 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
173 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGATTATTAGAAAAATAAACAATAG 114

Qy 61 GGGTTCCGCGCACATTTCCCGGAAAAAGTGCACCTGACGTC 101
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
113 GGGTTCCGCGCACATTTCCCGGAAAAAGTGCACCTGACGTC 73

RESULT 8
AAN81765/C
ID AAN81765 standard; DNA; 456 BP.
XX
AC AAN81765;
XX
XX 25-MAR-2003 (revised)
DT 13-DEC-1990 (first entry)

```

```

AAN70833/C
ID AAN70833 standard; DNA; 456 BP.
XX
AC AAN70833;
XX
XX 25-MAR-2003 (revised)
DT 10-MAR-2003 (revised)
DT 18-JAN-1991 (first entry)
XX
XX Beta-urogastrone sequence.
DE
XX Tumour; inosine; DNA probe; ds.
XX
XX Unidentified.
XX
XX Key Location/Qualifiers
FH promoter 125..170
FT /*tag= b
FT RBS 200..204
FT /*tag= c
FT CDS 209..439
FT /*tag= a
FT sig_peptide 209..277
FT /*tag= d
XX
XX JP62244398-A.
XX
XX 24-OCT-1987.
XX
XX 16-APR-1986; 86JP-00087368.
XX
XX 16-APR-1986; 86JP-00087368.
XX
XX (SEKI ) SEKISUI CHEM IND CO LTD.
XX
XX WPI; 1987-339045/48.
XX
XX P-PSDB; AAP70505.
XX
XX Detection of DNA and/or RNA - by converting to single strand form and
PT using probe contg. labelled inosine deriv.
XX
XX Disclosure; Page 11; 11pp; Japanese.
XX
XX An example of a sequence detected by a probe consisting of polyinosine,
CC polydeoxyinosine, oligoinosine and/or oligodeoxyinosine is labeled. The
CC ssDNA and probe are hybridized and the existence of DNA in the product is
CC detected. It can be used to detect the presence of malignant tumour.
CC (Updated on 10-MAR-2003 to add missing OS field.) (Updated on 25-MAR-2003
CC to correct PA field.)
XX
XX Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 101; DB 1; Length 456;
Best Local Similarity 100.0%; Pred. No. 2.6e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGATTATTAGAAAAATAAACAATAG 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
173 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGATTATTAGAAAAATAAACAATAG 114

Qy 61 GGGTTCCGCGCACATTTCCCGGAAAAAGTGCACCTGACGTC 101
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
113 GGGTTCCGCGCACATTTCCCGGAAAAAGTGCACCTGACGTC 73

RESULT 8
AAN81765/C
ID AAN81765 standard; DNA; 456 BP.
XX
AC AAN81765;
XX
XX 25-MAR-2003 (revised)
DT 13-DEC-1990 (first entry)

```

XX Sequence of HB101 (pGH201) encoding new beta-urogastrone deriv. Met (21),
DE Arg (53).
XX
XX Gastric acid secretion; cell proliferation; hormone; ds.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FH 209..277
CDS /*tag= a
XX 278..439
FT /*tag= b
FT /*tag= b
FT /*tag= b
FT /*tag= b
XX
XX JP63012298-A.
XX
XX 19-JAN-1988.
XX
XX 30-JUN-1986; 86JP-00153783.
XX
XX 30-JUN-1986; 86JP-00153783.
XX
XX (EART) EARTH SEIYAKU KK.
XX
XX WPI; 1988-054638/08.
DR P-PSDB; AAP81349.
XX
XX New beta-urogastrone deriv. - has gastric acid secretion inhibition and
PT proliferation promotion activity.
XX
XX Disclosure; Page 685; 76pp; Japanese.
XX
XX The deriv. has various biological activities such as gastric acid
CC secretion inhibiting action, or cell proliferation promoting action. The
CC deriv. has the same biological or pharmacological activities as beta-
CC urogastrone. It is not susceptible to denaturation by oxidn. and is
CC chemically stable. Deriv. has resistance to proteolytic enzymes such as
CC protease. (Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 456 BP; 115 A; 86 C; 104 G; 151 T; 0 U; 0 Other;
SQ
Query Match 100.0%; Score 101; DB 1; Length 456;
Best Local Similarity 100.0%; Pred. No. 2.6e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAATAAACAAATAG 60
DB 173 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAATAAACAAATAG 114
QY 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
DB 113 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 73
RESULT 9
ABA90413/c
ID ABA90413 standard; DNA; 466 BP.
XX
XX ABA90413;
XX
XX 12-FEB-2002 (first entry)
XX
XX Drosophila cell cycle progression protein coding sequence #48.
XX
XX Antiproliferative; cytostatic; cardiant; immunosuppressive; meiosis;
KW antiinflammatory; antipsoriatic; dermatological; antifungal; mitosis;
KW antiparasitic; antimalarial; antirheumatic; antiarthritic; cell division;
KW cell cycle progression protein; tumour; proliferative disorder;
KW cardiovascular; autoimmune; dermatological disorder; ds.
XX
XX Drosophila sp.
XX

PN WO200172774-A2.
XX
XX 04-OCT-2001.
XX
XX 23-MAR-2001; 2001WO-GB001297.
XX
XX 24-MAR-2000; 2000GB-00007268.
XX
XX (CYCL-) CYCLACEL LTD.
XX
XX Deak P, Glover DM, Midgley C;
PI WPI; 2002-055132/07.
DR
XX Polynucleotides encoding cell cycle progression proteins, useful for
PT treating a tumor or a proliferative disorder.
XX
XX Claim 1; Page 99; 213pp; English.
XX
XX The present invention relates to Drosophila cell cycle progression
CC proteins (AAM47572-AAM47608) and their coding sequences (ABA90366-
CC ABA90520). The coding sequences and proteins are useful for identifying a
CC substance capable of affecting the function of the corresponding gene, a
CC substance capable of inhibiting the cell division cycle, or capable of
CC inhibiting mitosis and/or meiosis. They can also be used in a method for
CC treating a tumour or proliferative disorder, cardiovascular disorders
CC (such as restenosis and cardiomyopathy), autoimmune disorders such as
CC (glomerulonephritis and rheumatoid arthritis), dermatological disorders
CC (such as psoriasis), antiinflammatory, antifungal and antiparasitic
CC disorders (such as malaria)
XX
XX Sequence 466 BP; 135 A; 87 C; 93 G; 150 T; 0 U; 1 Other;
SQ
Query Match 100.0%; Score 101; DB 6; Length 466;
Best Local Similarity 100.0%; Pred. No. 2.6e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAATAAACAAATAG 60
DB 280 AGGGTTATTGTCATGAGCGGATACATATTGTAATGTATTAGAAAATAAACAAATAG 221
QY 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
DB 220 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 180
RESULT 10
AAX21173/c
ID AAX21173 standard; DNA; 487 BP.
XX
XX AAX21173;
XX
XX 05-MAY-1999 (first entry)
XX
XX Polynucleotide sequence from the genome of Treponema pallidum.
XX
XX Treponema pallidum infection; syphilis; Borrelia infection; animal;
KW enzyme production; ds.
XX
XX Treponema pallidum.
XX
XX WO9859034-A2.
XX
XX 30-DEC-1998.
XX
XX 23-JUN-1998; 98WO-US013041.
XX
XX 24-JUN-1997; 97US-0050667P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Fraser CM;
XX

DR WPI; 1999-081273/07.
XX New isolated Treponema pallidum nucleic acids - used to develop products
PT for the detection, diagnosis, characterisation, prevention and therapy of
FT T. pallidum infections, particularly syphilis.
XX
XX Claim 1; Page 1106; 1150pp; English.
XX
CC AAX20500-21243 represent polynucleotide sequences from the genome of
CC Treponema pallidum. The sequences can be used for detection, diagnosis,
CC characterisation, prevention and therapy for T. pallidum infections,
CC particularly syphilis. They can also be used for detecting diseases
CC related to Borrelia infections in animals, and for the production of
CC biosynthetic products such as enzymes
XX
SQ Sequence 487 BP; 125 A; 127 C; 113 G; 121 T; 0 U; 1 Other;

Query Match 100.0%; Score 101; DB 2; Length 487;
Best Local Similarity 100.0%; Pred. No. 2.6e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 60
Db 323 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 264

Qy 61 GGGTTCCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101
Db 263 GGGTTCCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 223

RESULT 11
AAX21149/C
ID AAX21149 standard; DNA; 535 BP.
XX AC AAX21149;
XX
XX 05-MAY-1999 (first entry)
XX
XX Polynucleotide sequence from the genome of Treponema pallidum.
XX
XX Treponema pallidum infection; syphilis; Borrelia infection; animal;
KW enzyme production; ds.
XX
XX Treponema pallidum.
XX
XX WO9859034-A2.
XX
XX 30-DEC-1998.
XX
XX 23-JUN-1998; 98WO-US013041.
XX
XX 24-JUN-1997; 97US-0050667P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Fraser CM;
XX
XX WPI; 1999-081273/07.
XX
XX New isolated Treponema pallidum nucleic acids - used to develop products
PT for the detection, diagnosis, characterisation, prevention and therapy of
FT T. pallidum infections, particularly syphilis.
XX
XX Claim 1; Page 1093; 1150pp; English.
XX
CC AAX20500-21243 represent polynucleotide sequences from the genome of
CC Treponema pallidum. The sequences can be used for detection, diagnosis,
CC characterisation, prevention and therapy for T. pallidum infections,
CC particularly syphilis. They can also be used for detecting diseases
CC related to Borrelia infections in animals, and for the production of
CC biosynthetic products such as enzymes
XX
SQ Sequence 535 BP; 145 A; 108 C; 122 G; 155 T; 0 U; 5 Other;

Query Match 100.0%; Score 101; DB 2; Length 535;
Best Local Similarity 100.0%; Pred. No. 2.7e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 60
Db 158 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 99

Qy 61 GGGTTCCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 101
Db 98 GGGTTCCGCGCACATTTCCCGAAAGTGCCACCTGACGTC 58

RESULT 12
ABA90456/c
ID ABA90456 standard; DNA; 573 BP.
XX AC ABA90456;
XX
XX 12-FEB-2002 (first entry)
XX
XX Drosophila cell cycle progression protein coding sequence #91.
XX
XX Antiproliferative; cytostatic; cardiant; immunosuppressive; meiosis;
KW antiinflammatory; antipsoriatic; dermatological; antifungal; mitosis;
KW antiparasitic; antimalarial; antirheumatic; antiarthritic; cell division;
KW cell cycle progression protein; tumour; proliferative disorder;
KW cardiovascular; autoimmune; dermatological disorder; ds.
XX
XX Drosophila sp.
XX
XX WO200172774-A2.
XX
XX 04-OCT-2001.
XX
XX 23-MAR-2001; 2001WO-GB001297.
XX
XX 24-MAR-2000; 2000GB-00007268.
XX
XX (CYCL-) CYCLACEL LTD.
XX
XX Deak P, Glover DM, Midgley C;
XX
XX WPI; 2002-055132/07.
XX
XX Polynucleotides encoding cell cycle progression proteins, useful for
PT treating a tumor or a proliferative disorder.
XX
XX Claim 1; Page 144; 213pp; English.
XX
XX The present invention relates to Drosophila cell cycle progression
CC proteins (AAM47572-AAM47608) and their coding sequences (ABA90366-
CC ABA90520). The coding sequences and proteins are useful for identifying a
CC substance capable of affecting the function of the corresponding gene, a
CC substance capable of inhibiting the cell division cycle, or capable of
CC inhibiting mitosis and/or meiosis. They can also be used in a method for
CC treating a tumour or proliferative disorder, cardiovascular disorders
CC (such as restenosis and cardiomyopathy), autoimmune disorders such as
CC (glomerulonephritis and rheumatoid arthritis), dermatological disorders
CC (such as psoriasis), antiinflammatory, antifungal and antiparasitic
CC disorders (such as malaria)
XX
XX Sequence 573 BP; 154 A; 118 C; 116 G; 184 T; 0 U; 1 Other;

Query Match 100.0%; Score 101; DB 6; Length 573;
Best Local Similarity 100.0%; Pred. No. 2.7e-21;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 60
Db 355 AGGGTTATTGCTCATGAGCGGATACATATTGTAATGTATTAGAAAAATAAACAAATAG 296

QY 61 GGGTTCCGGCACATTTCCCGAAGAGTGCACCTGACGTC 101
|||||
Db 295 GGGTTCCGGCACATTTCCCGAAGAGTGCACCTGACGTC 255

RESULT 13

ADH58311

ID ADH58311 standard; DNA; 605 BP.

XX

AC ADH58311;

XX

XX 25-MAR-2004 (first entry)

XX

XX Electropherogram of a DNA sequencing reaction using E154 & T422.

DE

ds; primer library; extendable oligos; EO; ligation chain reaction; LCR;
KW rolling circle amplification; strand displacement amplification;
KW isothermal DNA amplification; biotechnology; agriculture;
KW medical research; pUC19 plasmid.

XX

OS Synthetic.

OS

Escherichia coli.

XX

PN WO2003093500-A1.

XX

PD 13-NOV-2003.

XX

XX 24-DEC-2002; 2002WO-AU001763.

XX

XX 01-MAY-2002; 2002AU-00002045.

XX

XX (NUCL-) NUCLEICS PTY LTD.

XX

XX Tillett D, Thomas T;

XX

XX WPI; 2004-053046/05.

XX

Increasing the affinity of an extendable oligonucleotide (EO) for a
target nucleic acid, for providing primers having improved specificity,
PT comprises hybridization of the EO to a template oligonucleotide (TO) and
extension of the EO.

XX

PS Example 10; Fig 23; 85pp; English.

XX

This invention relates to a novel method for the optimisation of primer
libraries. Specifically, it refers to increasing the affinity of short
oligonucleotide primers, also known as extendable oligos (EOs), for their
template sequences. The present invention describes improved methods for
sequencing and the linear and exponential amplification of DNA that can
be useful for PCR, RT-PCR, ligation chain reaction (LCR), rolling circle
amplification, strand displacement amplification and isothermal DNA
amplification. Accordingly, these extendable oligos with improved
specificity and affinity are particularly important in fields ranging
from biotechnology and agriculture to medical research. This
polynucleotide sequence is the electropherogram of a DNA sequencing
reaction that used the pUC19 plasmid and E154/T422 oligos, used in an
exemplification of the invention.

XX

SQ Sequence 605 BP; 159 A; 133 C; 147 G; 148 T; 0 U; 18 Other;

Query Match

Best Local Similarity 100.0%; Score 101; DB 12; Length 605;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGGTTATGTCTCATGACGGATACATATTGTAATGATTATTAGAAAATAACAAATAG 60

Db 259 AGGGTTATGTCTCATGACGGATACATATTGTAATGATTATTAGAAAATAACAAATAG 318

|||||

QY 61 GGGTTCCGGCACATTTCCCGAAGAGTGCACCTGACGTC 101

Db 319 GGGTTCCGGCACATTTCCCGAAGAGTGCACCTGACGTC 359

XX

RESULT 14

AAS30560/c

ID AAS30560 standard; DNA; 776 BP.

XX

AC AAS30560;

XX

XX 21-NOV-2001 (first entry)

XX

XX DNA encoding novel prostate gland antigen, Seq ID No 418.

XX

Human; neurotropic; neuroprotective; cytostatic; antiparkinsonian;
antianaeic; dermatological; immunosuppressive; antiinflammatory;
antiarthritic; antirheumatic; virucide; hepatotropic; nephrotropic;
osteopathic; prostate gland; prostatitis; adenocarcinoma; hair loss;
hyperplasia; carcinoma; adenocarcinoma; benign prostatic hypertrophy;
prostatic; prostate neoplastic disorder; skin aging;
reproductive system disorder; autoimmune disorder; urinary system;
systemic lupus erythematosus; rheumatoid arthritis; cardiovascular;
blood-related disorder; hyperproliferative disorder; respiratory;
neurological disorder; endocrine disorder; inflammatory disorder;
liver disorder; wound healing; food preservative; ds.

XX

OS Homo sapiens.

XX

XX WO200155447-A1.

XX

XX 02-AUG-2001.

XX

XX 17-JAN-2001; 2001WO-US001330.

XX

XX 31-JAN-2000; 2000US-0179065P.

XX

XX 04-FEB-2000; 2000US-0180628P.

XX

XX 24-FEB-2000; 2000US-0184664P.

XX

XX 02-MAR-2000; 2000US-0186350P.

XX

XX 16-MAR-2000; 2000US-0189874P.

XX

XX 17-MAR-2000; 2000US-0190076P.

XX

XX 18-APR-2000; 2000US-0198123P.

XX

XX 19-MAY-2000; 2000US-0205151P.

XX

XX 28-JUN-2000; 2000US-0214886P.

XX

XX 30-JUN-2000; 2000US-0215135P.

XX

XX 07-JUL-2000; 2000US-0216647P.

XX

XX 11-JUL-2000; 2000US-0217487P.

XX

XX 11-JUL-2000; 2000US-0217496P.

XX

XX 14-JUL-2000; 2000US-0218290P.

XX

XX 26-JUL-2000; 2000US-0220963P.

XX

XX 26-JUL-2000; 2000US-0220964P.

XX

XX 14-AUG-2000; 2000US-0224518P.

XX

XX 14-AUG-2000; 2000US-0224519P.

XX

XX 14-AUG-2000; 2000US-0225213P.

XX

XX 14-AUG-2000; 2000US-0225214P.

XX

XX 14-AUG-2000; 2000US-0225266P.

XX

XX 14-AUG-2000; 2000US-0225267P.

XX

XX 14-AUG-2000; 2000US-0225268P.

XX

XX 14-AUG-2000; 2000US-0225270P.

XX

XX 14-AUG-2000; 2000US-0225447P.

XX

XX 14-AUG-2000; 2000US-0225757P.

XX

XX 14-AUG-2000; 2000US-0225758P.

XX

XX 18-AUG-2000; 2000US-0225759P.

XX

XX 22-AUG-2000; 2000US-0226279P.

XX

XX 22-AUG-2000; 2000US-0226681P.

XX

XX 22-AUG-2000; 2000US-0226686P.

XX

XX 22-AUG-2000; 2000US-0227182P.

XX

XX 23-AUG-2000; 2000US-0227009P.

XX

XX 30-AUG-2000; 2000US-0228924P.

XX

XX 01-SEP-2000; 2000US-0229287P.

XX

XX 01-SEP-2000; 2000US-0229343P.

XX

XX 01-SEP-2000; 2000US-0229344P.

XX

XX 01-SEP-2000; 2000US-0229345P.

XX

XX 05-SEP-2000; 2000US-0229509P.

XX

XX 05-SEP-2000; 2000US-0229513P.

XX

XX 06-SEP-2000; 2000US-0230437P.

XX

XX AAS27819;
XX 07-NOV-2001 (first entry)
XX DNA encoding novel signal transduction pathway protein, Seq ID 1479.
XX Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
XX anti-inflammatory; anti-HIV; antibacterial; antineoplastic; cancer;
XX immune system disorder; rheumatoid arthritis; inflammatory condition;
XX organ transplant rejection; infection; hepatitis C; blood disorder;
XX sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
XX neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
XX chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
XX cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
XX reproductive system; gastrointestinal; liver disorder; AIDS; ds;
XX acquired immune deficiency syndrome.
OS Homo sapiens.
XX
XX WO200154733-A1.
XX
XX PD 02-AUG-2001.
XX
XX PF 17-JAN-2001; 2001WO-US001312.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225213P.
XX 14-AUG-2000; 2000US-0225214P.
XX 14-AUG-2000; 2000US-0225266P.
XX 14-AUG-2000; 2000US-0225267P.
XX 14-AUG-2000; 2000US-0225268P.
XX 14-AUG-2000; 2000US-0225270P.
XX 14-AUG-2000; 2000US-0225447P.
XX 14-AUG-2000; 2000US-0225757P.
XX 14-AUG-2000; 2000US-0225758P.
XX 14-AUG-2000; 2000US-0225759P.
XX 18-AUG-2000; 2000US-0226273P.
XX 22-AUG-2000; 2000US-0226681P.
XX 22-AUG-2000; 2000US-0226686P.
XX 23-AUG-2000; 2000US-0227182P.
XX 23-AUG-2000; 2000US-0227009P.
XX 30-AUG-2000; 2000US-0228924P.
XX 01-SEP-2000; 2000US-0228287P.
XX 01-SEP-2000; 2000US-0229343P.
XX 01-SEP-2000; 2000US-0229344P.
XX 01-SEP-2000; 2000US-0229345P.
XX 05-SEP-2000; 2000US-0229509P.
XX 05-SEP-2000; 2000US-0229513P.
XX 06-SEP-2000; 2000US-0230437P.
XX 06-SEP-2000; 2000US-0230438P.
XX 08-SEP-2000; 2000US-0231242P.
XX 08-SEP-2000; 2000US-0231243P.
XX 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 14, 2005, 05:15:57 ; Search time 961.667 Seconds
(without alignments)
3997.736 Million cell updates/sec

Title: US-09-482-682-8_COPY_7369_7469

Perfect score: 101

Sequence: 1 aggggtattgtctcatgagc.....gaaagtgcacctgacgtc 101

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_ges1:*
9: gb_ges2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	101	100.0	300	4	BM078095 83374 Heb
C 2	101	100.0	300	5	BU963956 EST88 Heb
C 3	101	100.0	300	5	BU964094 EST226 He
C 4	101	100.0	309	9	FR0009140
C 5	101	100.0	391	1	AL597149 DKFZp313J
C 6	101	100.0	414	9	CC819240 100005D19
C 7	101	100.0	417	4	BU684174
C 8	101	100.0	491	9	CC819923
C 9	101	100.0	495	4	BI805285
C 10	101	100.0	495	9	CC818374
C 11	101	100.0	496	9	CC818523
C 12	101	100.0	503	9	CC819854
C 13	101	100.0	515	9	CC817752
C 14	101	100.0	518	9	CC817128
C 15	101	100.0	519	9	CC817162
C 16	101	100.0	519	9	CC817796
C 17	101	100.0	521	9	CC819067
C 18	101	100.0	533	9	CC819841
C 19	101	100.0	542	9	CC816892
C 20	101	100.0	550	7	CR766622 DKFZp469H
C 21	101	100.0	551	9	CC816905
C 22	101	100.0	554	9	CC819058
C 23	101	100.0	563	9	CC819270
C 24	101	100.0	566	9	CC816848 100002D02

C 25	101	100.0	566	9	CC820024
C 26	101	100.0	567	9	CC817070
C 27	101	100.0	568	9	CC818640
C 28	101	100.0	571	9	CC816954
C 29	101	100.0	571	9	CC818423
C 30	101	100.0	580	9	CC819098
C 31	101	100.0	582	1	AL694813 DKFZp313I
C 32	101	100.0	583	9	CC817633
C 33	101	100.0	583	9	CC818436
C 34	101	100.0	586	9	CC816883
C 35	101	100.0	588	9	CC817788
C 36	101	100.0	588	9	CC818340
C 37	101	100.0	589	9	CC817595
C 38	101	100.0	590	9	CC819754
C 39	101	100.0	592	9	CC817679
C 40	101	100.0	592	9	CC818508
C 41	101	100.0	593	9	CC816942
C 42	101	100.0	593	9	CC817699
C 43	101	100.0	594	9	CC818287
C 44	101	100.0	594	9	CC818422
C 45	101	100.0	595	9	CC816929

ALIGNMENTS

RESULT 1
BM078095/c 300 bp mRNA linear EST 30-NOV-2001
LOCUS 83374 Hebeloma cylindrosporum functional cDNA library Hebeloma
DEFINITION BM078095
ACCESSION BM078095
VERSION BM078095.1 GI:17157967
KEYWORDS EST.
SOURCE Hebeloma cylindrosporum
ORGANISM Hebeloma cylindrosporum
REFERENCE Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
Agaricales; Cortinariaceae; Hebeloma.
AUTHORS Wipf, D., Benjdia, M., Tegeder, M. and Frommer, W.B.
TITLE Construction of a functional cDNA library from the ectomycorrhizal fungus Hebeloma cylindrosporum
JOURNAL Unpublished (2001)
COMMENT Contact: Wipf D
ZMBP - Center for Molecular Biology of Plants
University of Tuebingen
Auf der Morgenstelle 1, 72076 Tuebingen, Germany
Tel: 49 7071 2976160
Fax: 49 7071 293287
Email: daniel.wipf@zmbp.uni-tuebingen.de
PCR Primers
FORWARD: pDR196 5' primer (PMA 5')
High quality sequence stop: 300
POLYA-No.

FEATURES

Location/Qualifiers
1..300
/organism="Hebeloma cylindrosporum"
/mol_type="mRNA"
/strain="H1"
/db_xref="taxon:76867"
/tissue_type="Mycelia"
/lab_host="E. coli XL1-Blue"
/clone_lib="Hebeloma cylindrosporum functional cDNA library"
/note="Vector: pDR 196 (unpublished); Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Query Match 100.0%; Score 101; DB 4; Length 300;
Best Local Similarity 100.0%; Pred. No. 8.1e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AGGTTATTGCTCATGCGGATACATATTGTAATGTTAGAAAATAACAAATAG 60

```

|||||
174 AGGTTATTGTCATGAGCGGATACATATTGTAATGATTATTAGAAAATAAACAAATAG 115
|||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGAGTC 101
|||||
Db 114 GGGTTCGCGCACATTTCCCGAAAGTGCACCTGAGTC 74
|||||

RESULT 2
BU963956/c
LOCUS
DEFINITION EST88 Hebeloma cylindrosporium functional cDNA library Hebeloma
cylindrosporium cDNA, mRNA sequence.
ACCESSION BU963956
VERSION BU963956.1 GI:24204753
KEYWORDS
SOURCE
ORGANISM Hebeloma cylindrosporium
Hebeloma cylindrosporium
Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
Agaricales; Cortinariaceae; Hebeloma.
REFERENCE
AUTHORS Wipf,D., Benjdia,M., Rikirsch,E., Zimmermann,S., Tegeder,M. and
Frommer,W.B.
TITLE A Hebeloma cylindrosporium expression cDNA library for suppression
cloning in yeast mutants, complementation of a yeast his4 mutant
and EST analysis
JOURNAL
COMMENT Unpublished (2002)
Contact: Wipf D
ZMBP - Center for Molecular Biology of Plants
University of Tuebingen
Auf der Morgenstelle 1, 72076 Tuebingen, Germany
Tel: 49 7071 2976160
Fax: 49 7071 293287
Email: daniel.wipf@zmbp.uni-tuebingen.de
homolog to Enterobacteria phage f1 ; ampicillinase (1e-10).

FEATURES
source
1..300
/organism="Hebeloma cylindrosporium"
/mol_type="mRNA"
/db_xref="taxon:76867"
/tissue_type="Mycelia"
/lab_host="E. coli XL1-Blue"
/clone_lib="Hebeloma cylindrosporium functional cDNA
library"
/notes="vector: pDR 196 (unpublished); Site_1: EcoRI;
Site_2: XhoI"

ORIGIN
Query Match 100.0%; Score 101; DB 5; Length 300;
Best Local Similarity 100.0%; Pred. No. 8.1e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGATTATTAGAAAATAAACAAATAG 60
|||||
Db 170 AGGGTTATTGTCATGAGCGGATACATATTGTAATGATTATTAGAAAATAAACAAATAG 111
|||||

RESULT 3
BU964094/c
LOCUS
DEFINITION EST226 Hebeloma cylindrosporium functional cDNA library Hebeloma
cylindrosporium cDNA, mRNA sequence.
ACCESSION BU964094
VERSION BU964094.1 GI:24204891
KEYWORDS
SOURCE
ORGANISM Hebeloma cylindrosporium
Hebeloma cylindrosporium
Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
Agaricales; Cortinariaceae; Hebeloma.
REFERENCE
AUTHORS Wipf,D., Benjdia,M., Rikirsch,E., Zimmermann,S., Tegeder,M. and
Frommer,W.B.
TITLE A Hebeloma cylindrosporium expression cDNA library for suppression
cloning in yeast mutants, complementation of a yeast his4 mutant
and EST analysis
JOURNAL
COMMENT Unpublished (2002)
Contact: Wipf D
ZMBP - Center for Molecular Biology of Plants
University of Tuebingen
Auf der Morgenstelle 1, 72076 Tuebingen, Germany
Tel: 49 7071 2976160
Fax: 49 7071 293287
Email: daniel.wipf@zmbp.uni-tuebingen.de
homolog to Enterobacteria phage f1 ; ampicillinase (1e-10).

FEATURES
source
1..300
/organism="Hebeloma cylindrosporium"
/mol_type="mRNA"
/db_xref="taxon:76867"
/tissue_type="Mycelia"
/lab_host="E. coli XL1-Blue"
/clone_lib="Hebeloma cylindrosporium functional cDNA
library"
/notes="vector: pDR 196 (unpublished); Site_1: EcoRI;
Site_2: XhoI"

ORIGIN
Query Match 100.0%; Score 101; DB 5; Length 300;
Best Local Similarity 100.0%; Pred. No. 8.1e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGATTATTAGAAAATAAACAAATAG 60
|||||
Db 170 AGGGTTATTGTCATGAGCGGATACATATTGTAATGATTATTAGAAAATAAACAAATAG 111
|||||

RESULT 4
FR0009140
LOCUS
DEFINITION F.rubripes GSS sequence, clone 010H20aC4, genomic survey sequence.
ACCESSION AL000426
VERSION AL000426.1 GI:2438278
KEYWORDS GSS; genome survey sequence.
SOURCE Takifugu rubripes (Fugu rubripes)
ORGANISM Takifugu rubripes
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Takifugu.

REFERENCE
AUTHORS Elgar,G., Clark,M.S., Meek,S., Smith,S., Warner,S., Edwards,Y.J.,
Bouchireb,N., Cottage,A., Yeo,G.S., Umrantia,Y., Williams,G. and
Brenner,S.
TITLE Generation and analysis of 25 Mb of genomic DNA from the pufferfish
Fugu rubripes by sequence scanning
JOURNAL Genome Res. 9 (10), 960-971 (1999)
MEDLINE 99455097
PUBMED 10523524
REFERENCE
AUTHORS Elgar,G., Clark,M., Smith,S., Meek,S., Warner,S., Umrantia,Y.,
Williams,G. and Brenner,S.
TITLE Direct Submission
JOURNAL Submitted (09-SEP-1997) MRC Human Genome Mapping Project Resource
Centre Hinxton, Cambridge, CB10 1SB. Email: biohelp@hgmrc.ac.uk
COMMENT Vector: pBluescript II KS
V_type: phagemid

```

```

Agaricales; Cortinariaceae; Hebeloma.
1 (bases 1 to 300)
Wipf,D., Benjdia,M., Rikirsch,E., Zimmermann,S., Tegeder,M. and
Frommer,W.B.
TITLE A Hebeloma cylindrosporium expression cDNA library for suppression
cloning in yeast mutants, complementation of a yeast his4 mutant
and EST analysis
JOURNAL
COMMENT Unpublished (2002)
Contact: Wipf D
ZMBP - Center for Molecular Biology of Plants
University of Tuebingen
Auf der Morgenstelle 1, 72076 Tuebingen, Germany
Tel: 49 7071 2976160
Fax: 49 7071 293287
Email: daniel.wipf@zmbp.uni-tuebingen.de
no homology below 1e-10.

FEATURES
Location/Qualifiers
1..300
/organism="Hebeloma cylindrosporium"
/mol_type="mRNA"
/db_xref="taxon:76867"
/tissue_type="Mycelia"
/lab_host="E. coli XL1-Blue"
/clone_lib="Hebeloma cylindrosporium functional cDNA
library"
/notes="vector: pDR 196 (unpublished); Site_1: EcoRI;
Site_2: XhoI"

ORIGIN
Query Match 100.0%; Score 101; DB 5; Length 300;
Best Local Similarity 100.0%; Pred. No. 8.1e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATTGTCATGAGCGGATACATATTGTAATGATTATTAGAAAATAAACAAATAG 60
|||||
Db 170 AGGGTTATTGTCATGAGCGGATACATATTGTAATGATTATTAGAAAATAAACAAATAG 111
|||||

RESULT 4
FR0009140
LOCUS
DEFINITION F.rubripes GSS sequence, clone 010H20aC4, genomic survey sequence.
ACCESSION AL000426
VERSION AL000426.1 GI:2438278
KEYWORDS GSS; genome survey sequence.
SOURCE Takifugu rubripes (Fugu rubripes)
ORGANISM Takifugu rubripes
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae; Takifugu.

REFERENCE
AUTHORS Elgar,G., Clark,M.S., Meek,S., Smith,S., Warner,S., Edwards,Y.J.,
Bouchireb,N., Cottage,A., Yeo,G.S., Umrantia,Y., Williams,G. and
Brenner,S.
TITLE Generation and analysis of 25 Mb of genomic DNA from the pufferfish
Fugu rubripes by sequence scanning
JOURNAL Genome Res. 9 (10), 960-971 (1999)
MEDLINE 99455097
PUBMED 10523524
REFERENCE
AUTHORS Elgar,G., Clark,M., Smith,S., Meek,S., Warner,S., Umrantia,Y.,
Williams,G. and Brenner,S.
TITLE Direct Submission
JOURNAL Submitted (09-SEP-1997) MRC Human Genome Mapping Project Resource
Centre Hinxton, Cambridge, CB10 1SB. Email: biohelp@hgmrc.ac.uk
COMMENT Vector: pBluescript II KS
V_type: phagemid

```

```

PRIMER: KS
DESCR:
One pass dye-terminator sequencing of cosmid cloned genomic
sequence.
FEATURES
    source
        Location/Qualifiers
            1..309
                /organism="Takifugu rubripes"
                /mol_type="genomic DNA"
                /db_xref="taxon:31033"
                /clone="010H20aC4"
                /clone_lib="cosmid 010H20"
ORIGIN
    Query Match      100.0%; Score 101; DB 9; Length 309;
    Best Local Similarity 100.0%; Pred. No. 8.2e-19;
    Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAACAATAG 60
Db 39 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAACAATAG 98
QY 61 GGGTCCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 101
Db 99 GGGTCCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 139
RESULT 5
LOCUS AL597149 391 bp mRNA linear EST 04-SEP-2003
DEFINITION DKFP313J1611_r1 313 (synonym: hlcc2) Homo sapiens cDNA clone
ACCESSION AL597149
VERSION AL597149.1 GI:15154845
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Koehrer,K., Beyer,A., Mewes,W., Weil,B. and Wiemann,S.
TITLE EST (Koehrer K., Beyer A., Mewes H.W., Weil B. and Wiemann S.)
JOURNAL Unpublished (1999)
COMMENT Contact: MIPS
MIPS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by BMFZ (Biomedical Research Center at the Charite,
Berlin/Germany) within the cDNA sequencing consortium of the German
Genome Project.
No sl sequence available.
This clone (DKFP313J1611) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
    source
        Location/Qualifiers
            1..391
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="DKFP313J1611"
                /dev_stage="adult"
                /lab_host="DH10B"
                /clone_lib="313 (synonym: hlcc2)"
                /note="vector: pTriplex2; Site_1: sf1A; Site_2: sf1B;
                cDNA-collection"
ORIGIN
    Query Match      100.0%; Score 101; DB 1; Length 391;
    Best Local Similarity 100.0%; Pred. No. 8.3e-19;
    Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAACAATAG 60
Db 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAACAATAG 355
PRIMER: KS
DESCR:
One pass dye-terminator sequencing of cosmid cloned genomic
sequence.
FEATURES
    source
        Location/Qualifiers
            1..309
                /organism="Takifugu rubripes"
                /mol_type="genomic DNA"
                /db_xref="taxon:31033"
                /clone="010H20aC4"
                /clone_lib="cosmid 010H20"
ORIGIN
    Query Match      100.0%; Score 101; DB 9; Length 309;
    Best Local Similarity 100.0%; Pred. No. 8.2e-19;
    Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAACAATAG 60
Db 39 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAACAATAG 98
QY 61 GGGTCCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 101
Db 99 GGGTCCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 139
RESULT 6
LOCUS CC819240/c 414 bp DNA linear GSS 17-JUL-2003
DEFINITION CC819240
ACCESSION CC819240
VERSION CC819240.1 GI:32899308
KEYWORDS GSS.
SOURCE Sterkiella histriomuscorum (Oxytricha trifallax)
ORGANISM Sterkiella histriomuscorum
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
Stichotrichida; Oxytrichidae; Sterkiella.
REFERENCE
    1 (bases 1 to 414)
    Dunn,D., Doak,T., Herrick,G. and Weiss,R.
    Paired end reads from plasmid inserts of Oxytricha trifallax
    macronuclear chromosomes
    Unpublished (2003)
    Contact: Robert B. Weiss
    University of Utah Genome Center
    Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
    84112, USA
    Tel: 801 585 5606
    Fax: 801 585 7177
    Email: ddunn@genetics.utah.edu
    Plate: 0005 row: D column: 19
    Seq primer: CACACAGGAACAGCTATGACC
    Class: plasmid ends
    High quality sequence stop: 414.
    Location/Qualifiers
        1..414
            /organism="Sterkiella histriomuscorum"
            /mol_type="genomic DNA"
            /db_xref="taxon:94289"
            /clone="UUGC100005D19"
            /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
            /clone_lib="Oxytricha plasmid UUGC10 library"
            /note="vector: PWD42nv; Purified macronuclear chromosomal
            DNA from Oxytricha trifallax was blunt end-repaired with
            T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
            oligonucleotides were ligated to the blunt ends in high
            molar excess. Vector DNA was prepared from a derivative of
            PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
            derivative of plasmid RL. The vector was ligated with
            adaptors complementary to the insert adaptors and
            purified. The sheared, adaptor mouse DNA was annealed to
            adaptor vector DNA, and transformed into
            chemically-competent E. Coli XL10-Gold (Stratagene) cells
            and selected for ampicillin resistance."
ORIGIN
    Query Match      100.0%; Score 101; DB 9; Length 414;
    Best Local Similarity 100.0%; Pred. No. 8.3e-19;
    Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAACAATAG 60
Db 414 AGGGTTATTGCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAACAATAG 355
QY 61 GGGTCCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 101
Db 354 GGGTCCGCGCACATTTCCCGAAAAGTCCACCTGACGTC 314

```

```

RESULT 7
BJ684174/c
LOCUS BJ684174 HCEST library Haplochromis chilotes cDNA clone no90c12,
DEFINITION BJ684174 HCEST library Haplochromis chilotes cDNA clone no90c12,
RNA sequence.
ACCESSION BJ684174
VERSION BJ684174.1 GI:46527295
KEYWORDS EST.
SOURCE Haplochromis chilotes
ORGANISM Haplochromis chilotes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Acanthopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percormorpha; Perciformes;
Labroidae; Cichlidae; Haplochromis.
REFERENCE 1 (bases 1 to 417)
Watanabe,M., Kobayashi,N., Shin-I,T., Kohara,Y. and Okada,N.
Orf sequences of cichlid in Lake Victoria are essentially same
Unpublished (2004)
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@gene.nig.ac.jp.
FEATURES             source
1..417
/organism="Haplochromis chilotes"
/mol_type="mRNA"
/db_xref="taxon:257977"
/clone="no90c12"
/tissue_type="jaw"
/dev_stages="varied"
/clone_lib="HCEST library"

ORIGIN
Query Match      100.0%; Score 101; DB 4; Length 417;
Best Local Similarity 100.0%; Pred. No. 8.3e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AGGGTTATTGCTCATGCGGATACATATTGTAATGTTATTAGAAAATAAACAAATAG 60
Db 129 AGGGTTATTGCTCATGCGGATACATATTGTAATGTTATTAGAAAATAAACAAATAG 70

Oy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCACCTGAGCTC 101
Db 69 GGGTTCGCGCACATTTCCCGAAAAGTGCACCTGAGCTC 29

RESULT 8
CC819923/c
LOCUS CC819923
DEFINITION CC819923 Oxytricha plasmid UUGC10006J13 R, genomic survey
sequence.
ACCESSION CC819923
VERSION CC819923.1 GI:32900671
KEYWORDS GSS.
SOURCE Sterkiella histriomscorum (Oxytricha trifallax)
ORGANISM Sterkiella histriomscorum
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
Stichotrichida; Oxytrichidae; Sterkiella.
REFERENCE 1 (bases 1 to 491)
Dunn,D., Doak,T., Herrick,G. and Weiss,R.
Paired end reads from plasmid inserts of Oxytricha trifallax
macronuclear chromosomes
Unpublished (2003)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606

```

```

Fax: 801 585 7177
Email: dunn@genetics.utah.edu
Plate: 0006 row: J column: 13
Seq primer: CACACAGGAACACGCTATGACC
Class: plasmid ends
High quality sequence stop: 491.
FEATURES             source
1..491
/organism="Sterkiella histriomscorum"
/mol_type="genomic DNA"
/db_xref="taxon:94289"
/clone="UUGC10006J13"
/lab_hosts="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Oxytricha plasmid UUGC10 library"
/notes="Vector: PWD42nv; Purified macronuclear chromosomal
DNA from Oxytricha trifallax was blunt end-repaired with
T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
oligonucleotides were ligated to the blunt ends in high
molar excess. Vector DNA was prepared from a derivative of
pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
derivative of plasmid R1. The vector was ligated with
adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. Coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

ORIGIN
Query Match      100.0%; Score 101; DB 9; Length 491;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 AGGGTTATTGCTCATGCGGATACATATTGTAATGTTATTAGAAAATAAACAAATAG 60
Db 412 AGGGTTATTGCTCATGCGGATACATATTGTAATGTTATTAGAAAATAAACAAATAG 353

Oy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCACCTGAGCTC 101
Db 352 GGGTTCGCGCACATTTCCCGAAAAGTGCACCTGAGCTC 312

RESULT 9
BI805285
LOCUS BI805285
DEFINITION BI805285 Stem library from Oryza sativa (3-5 leaf stage) Oryza
sativa cDNA clone S035A01, mRNA sequence.
ACCESSION BI805285
VERSION BI805285.1 GI:15852489
KEYWORDS EST.
SOURCE Oryza sativa
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartidae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 495)
Dong,H.T., Li,D.B., Zhuang,X.F., Dai,C.G., Sun,L.X., Pei,Y.X.,
Wu,H.F., Jiang,Y.X., Yu,F.C., Gao,Q.K. and Lou,Y.C.
A Gene Expression Screen in Oryza sativa
Unpublished (2001)
Contact: Haitao Dong, Debao Li
Bioinformatics and Gene Network Research Group
Zhejiang University
Kaixuan Road 268#, Hangzhou, Zhejiang, P.R.China
Tel: 0086-571-86892051
Fax: 0086-571-86961525
Email: webmaster@estarray.org, URL: http://www.estarray.org
Seq primer: M13 forward primer.
FEATURES             source
1..495
/organism="Oryza sativa"
/mol_type="mRNA"
/db_xref="taxon:4530"
/clone="S035A01"

```

/tissue_type="Stem"
/dev_stages="3-5 leaf stage"
/clone_lib="Stem library from Oryza sativa (3-5 leaf stage)"
/note="Vector: pSport2"

ORIGIN

Query Match 100.0%; Score 101; DB 4; Length 495;
Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATGTCATGAGCGGATACATATTGATGTATTAGAAAATAAACAATAG 60
|||||
Db 62 AGGGTTATGTCATGAGCGGATACATATTGATGTATTAGAAAATAAACAATAG 121
|||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
|||||
Db 122 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 162
|||||

RESULT 10

CC818374/c
LOCUS CC818374.1 GI:32897661 495 bp DNA linear GSS 17-JUL-2003
DEFINITION 100004807R Oxytricha plasmid UUGC10004B07 R, genomic survey
histriomuscorum sequence.

ACCESSION

VERSION CC818374

KEYWORDS

SOURCE GSS

ORGANISM

Sterkiella histriomuscorum (Oxytricha trifallax)
Sterkiella histriomuscorum
Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;
Stichotrichida; Oxytrichidae; Sterkiella.
1 (bases 1 to 495)
Dunn, D., Doak, T., Herrick, G. and Weiss, R.
Paired end reads from plasmid inserts of Oxytricha trifallax
macronuclear chromosomes

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Plate: 0004 row: B column: 07
Seq primer: CACACAGGAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 495.

FEATURES

source

1..495
/organism="Sterkiella histriomuscorum"
/mol_type="genomic DNA"
/db_xref="taxon:94289"
/clone="UUGC10004B07"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/clone_lib="Oxytricha plasmid UUGC10 library"
/note="Vector: PWD42nv; Purified macronuclear chromosomal
DNA from Oxytricha trifallax was blunt end-repaired with
T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
oligonucleotides were ligated to the blunt ends in high
molar excess. Vector DNA was prepared from a derivative of
pWD42 [gi|4732114|gb|AF129072.1], a copy-number inducible
derivative of plasmid R1. The vector was ligated with
adaptors complementary to the insert adaptors and
purified. The sheared, adaptor mouse DNA was annealed to
adaptor vector DNA, and transformed into
chemically-competent E. Coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

ORIGIN

Query Match 100.0%; Score 101; DB 9; Length 495;

Best Local Similarity 100.0%; Pred. No. 8.4e-19;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATGTCATGAGCGGATACATATTGATGTATTAGAAAATAAACAATAG 60
|||||
Db 392 AGGGTTATGTCATGAGCGGATACATATTGATGTATTAGAAAATAAACAATAG 333
|||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
|||||
Db 332 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 292
|||||

RESULT 11

CC818523/c

LOCUS CC818523.1 GI:32897943 496 bp DNA linear GSS 17-JUL-2003

DEFINITION 100004L13R Oxytricha plasmid UUGC10 library Sterkiella

histriomuscorum genomic clone UUGC100004L13 R, genomic survey

sequence.

ACCESSION CC818523

VERSION CC818523.1 GI:32897943

KEYWORDS GSS

SOURCE Sterkiella histriomuscorum (Oxytricha trifallax)

ORGANISM Sterkiella histriomuscorum

Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichia;

Stichotrichida; Oxytrichidae; Sterkiella.

1 (bases 1 to 496)

Dunn, D., Doak, T., Herrick, G. and Weiss, R.

Paired end reads from plasmid inserts of Oxytricha trifallax

macronuclear chromosomes

Unpublished (2003)

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Plate: 0004 row: L column: 13

Seq primer: CACACAGGAACAGCTATGACC

Class: plasmid ends

High quality sequence stop: 496.

Location/Qualifiers

1..496

/organism="Sterkiella histriomuscorum"

/mol_type="genomic DNA"

/db_xref="taxon:94289"

/clone="UUGC100004L13"

/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"

/clone_lib="Oxytricha plasmid UUGC10 library"

/note="Vector: PWD42nv; Purified macronuclear chromosomal

DNA from Oxytricha trifallax was blunt end-repaired with

T4 DNA polymerase and T4 polynucleotide kinase. Adaptor

oligonucleotides were ligated to the blunt ends in high

molar excess. Vector DNA was prepared from a derivative of

pWD42 [gi|4732114|gb|AF129072.1], a copy-number inducible

derivative of plasmid R1. The vector was ligated with

adaptors complementary to the insert adaptors and

purified. The sheared, adaptor mouse DNA was annealed to

adaptor vector DNA, and transformed into

chemically-competent E. Coli XL10-Gold (Stratagene) cells

and selected for ampicillin resistance."

ORIGIN

Query Match 100.0%; Score 101; DB 9; Length 496;

Best Local Similarity 100.0%; Pred. No. 8.4e-19;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGGGTTATGTCATGAGCGGATACATATTGATGTATTAGAAAATAAACAATAG 60
|||||
Db 391 AGGGTTATGTCATGAGCGGATACATATTGATGTATTAGAAAATAAACAATAG 332
|||||

Qy 61 GGGTTCGCGCACATTTCCCGAAAAGTGCCACCTGACGTC 101
|||||

RESULT 13

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Stichotrichida; Oxytrichidae; Sterkiella.
 1 (bases 1 to 518)
 Dunn, D., Doak, T., Herrick, G. and Weiss, R.
 Paired end reads from plasmid inserts of Oxytricha trifallax
 macronuclear chromosomes
 Unpublished (2003)
 Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Plate: 0002 row: D column: 21
 Seq primer: CACACAGGAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 518.

FEATURES
 source
 1..518
 Location/Qualifiers
 /organism="Sterkiella histriomuscorum"
 /mol_type="genomic DNA"
 /db_xref="taxon:94289"
 /clone="UUGC100002D21"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /clone_lib="Oxytricha plasmid UUGC10 library"
 /note="Vector: PWD42nv; Purified macronuclear chromosomal
 DNA from Oxytricha trifallax was blunt end-repaired with
 T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
 oligonucleotides were ligated to the blunt ends in high
 molar excess. Vector DNA was prepared from a derivative of
 pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
 derivative of plasmid R1. The vector was ligated with
 adaptors complementary to the insert adaptors and
 purified. The sheared, adapted mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. Coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

ORIGIN
 Query Match 100.0%; Score 101; DB 9; Length 518;
 Best Local Similarity 100.0%; Pred. No. 8.4e-19;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 AGGGTTATTGCTCATGCGGATACATATTTGAATGTATTAGAAAATAAACAAATAG 60
 Db 410 AGGGTTATTGCTCATGCGGATACATATTTGAATGTATTAGAAAATAAACAAATAG 351
 Qy 61 GGGTTCCGCGACATTTCCCGAAAAGTGCCACCTGACGTC 101
 Db 350 GGGTTCCGCGACATTTCCCGAAAAGTGCCACCTGACGTC 310
 RESULT 15
 CC817162/c
 LOCUS
 DEFINITION
 histriomuscorum genomic clone UUGC100002J19 R, genomic survey
 sequence.
 CC817162
 CC817162.1 GI:32896449
 GSS
 Sterkiella histriomuscorum (Oxytricha trifallax)
 Sterkiella histriomuscorum
 Eukaryota; Alveolata; Ciliophora; Spirotrichea; Stichotrichida;
 Stichotrichida; Oxytrichidae; Sterkiella.
 1 (bases 1 to 519)
 Dunn, D., Doak, T., Herrick, G. and Weiss, R.
 Paired end reads from plasmid inserts of Oxytricha trifallax
 macronuclear chromosomes
 Unpublished (2003)
 Contact: Robert B. Weiss
 University of Utah Genome Center

University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Plate: 0002 row: J column: 19
 Seq primer: CACACAGGAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 519.

FEATURES
 source
 1..519
 Location/Qualifiers
 /organism="Sterkiella histriomuscorum"
 /mol_type="genomic DNA"
 /db_xref="taxon:94289"
 /clone="UUGC100002J19"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /clone_lib="Oxytricha plasmid UUGC10 library"
 /note="Vector: PWD42nv; Purified macronuclear chromosomal
 DNA from Oxytricha trifallax was blunt end-repaired with
 T4 DNA polymerase and T4 polynucleotide kinase. Adaptor
 oligonucleotides were ligated to the blunt ends in high
 molar excess. Vector DNA was prepared from a derivative of
 pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible
 derivative of plasmid R1. The vector was ligated with
 adaptors complementary to the insert adaptors and
 purified. The sheared, adapted mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. Coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

ORIGIN
 Query Match 100.0%; Score 101; DB 9; Length 519;
 Best Local Similarity 100.0%; Pred. No. 8.4e-19;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 AGGGTTATTGCTCATGCGGATACATATTTGAATGTATTAGAAAATAAACAAATAG 60
 Db 416 AGGGTTATTGCTCATGCGGATACATATTTGAATGTATTAGAAAATAAACAAATAG 357
 Qy 61 GGGTTCCGCGACATTTCCCGAAAAGTGCCACCTGACGTC 101
 Db 356 GGGTTCCGCGACATTTCCCGAAAAGTGCCACCTGACGTC 316
 Search completed: July 14, 2005, 23:22:32
 Job time : 963.667 secs

THIS PAGE BLANK (USPTO)